INNOVATIONS IN AUDIO • AUDIO ELECTRONICS • THE BEST IN DIY AUDIO

www.audioxpress.com

audioxpress

ADVANCING THE EVOLUTION OF AUDIO TECHNOLOGY

Market Update

True Wireless Stereo Earbuds

The Next Frontier in Voice Technology

By J. Martins

Hearables in Focus

Putting the Smart in Smart Devices: The Future of Hearables Is Voice Controlled

By Probal Lala (Fluent.ai)

The Acoustical Challenges of Designing True Wireless Earphones

By Andrew Bellavia (Knowles)

You Can DIY!

Building a Guitar-Controlled Synthesizer: The Sample & Hold Time Machine

By Ethan Winer

Speakers

The Evolution of Production Line Rub & Buzz Measurements

By Steve Temme (Listen, Inc.)

Sound Control

Frequency Response of a Listening Room

By Richard Honeycutt

Audio Electronics

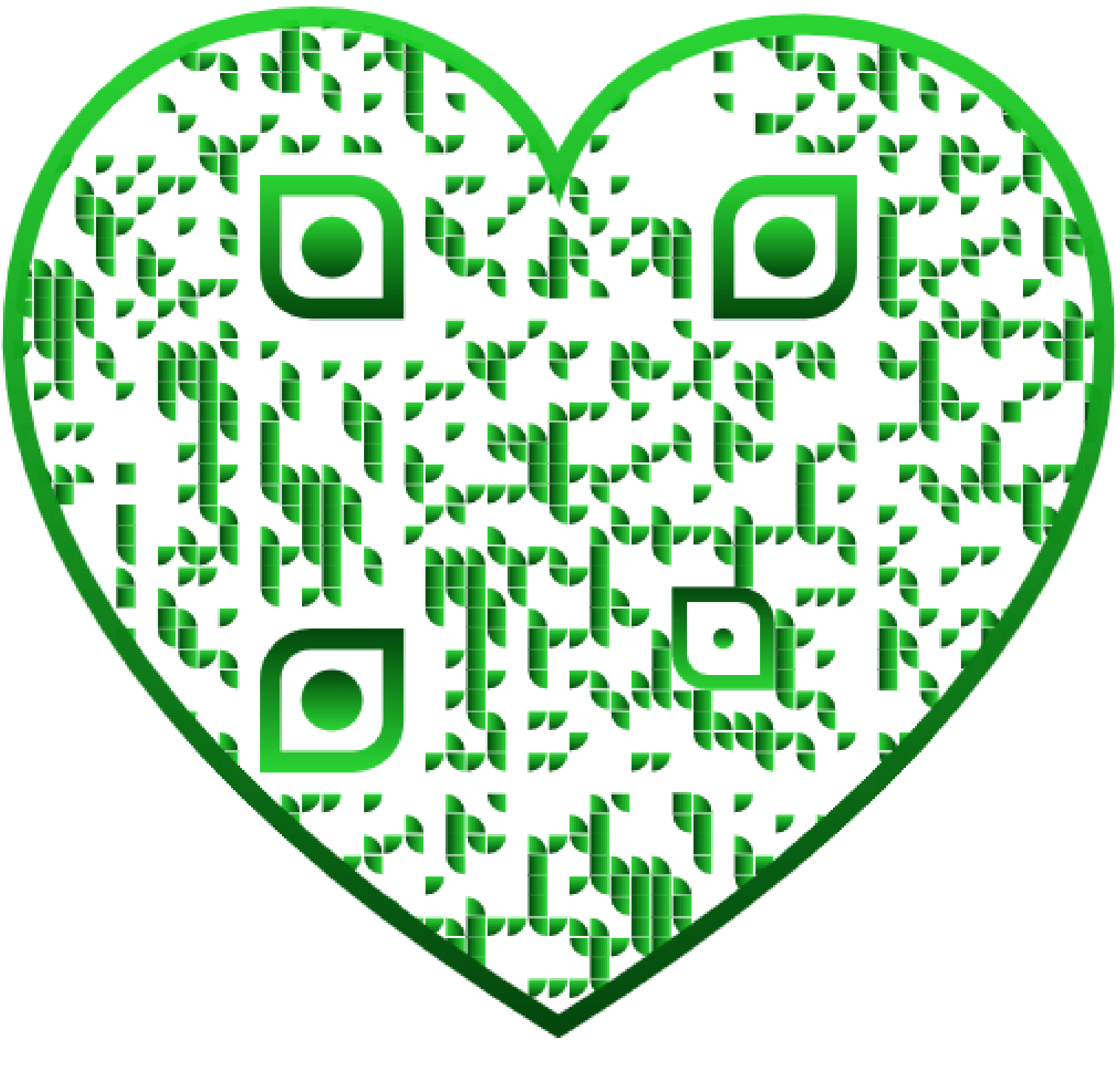
GaN Technology in Audio Power Amplification

By Stuart Yaniger

$9.00US $10.00CAN

04

APRIL 2022



NEW

Western Electric
MAKER OF ELECTRON TUBES AND HIGH FIDELITY

91E Integrated Amplifier

Available in nickel, gold, and black.
Matched pair Western Electric 300B electron tubes included.

THE FINEST INTEGRATED EVER MADE

Introducing the 91E, a new integrated amplifier from Western Electric. Meet ambitious, single-ended (SE) sound inspired by the 91A, its famous predecessor in high-fidelity for early Hollywood motion pictures. A groundbreaking new SE design evokes a classic feel with a sonic quality and resolution fit for today's listeners.

Experience over 20 watts per channel—a first for the 300B. Developed over two years, the 91E's patented circuitry intelligently steers the parallel feed constant current source, allowing the 300B to deliver more than double its previously rated output of 7-8 watts. You've never heard the 300B like this before.

FOR PRICING & AVAILABILITY VISIT

westernelectric.com/91e

April 2022 ISSN 1548-0628

www.audioxpress.com

audioXpress (US ISSN 1548-0628) is published monthly, at $50 per year for the US, at $65 per year for Canada, and at $75 per year Foreign/ROW, by KCK Media Corp., P.O. Box 417, Chase City, VA 23924, US Periodical Postage paid at Chase City, VA, and additional offices.

Head Office:
KCK Media Corp.
P.O. Box 417
Chase City, VA 23924, US
Phone: 434-533-0246
Fax: 888-980-1303

Subscription Management:
audioXpress
P.O. Box 417
Chase City, VA 23924, US
Phone: 434-533-0246
E-mail: custserv@audioxpress.com
Internet: www.audioxpress.com

Postmaster: Send address changes to:
audioXpress
P.O. Box 417
Chase City, VA 23924, US

Advertising:
Strategic Media Marketing, Inc.
2 Main Street
Gloucester, MA 01930, US
Phone: 978-281-7708
Fax: 978-281-7706
E-mail: audioxpress@smmarketing.us
Advertising rates and terms available on request.

Editorial Inquiries:
Send editorial correspondence, suggestions, and contributed articles to:
audioXpress, Editorial Department
P.O. Box 417
Chase City, VA 23924, US
E-mail: editor@audioxpress.com

Legal Notice:
Each design published in audioXpress is the intellectual property of its author and is offered to readers for their personal use only.
Any commercial use of such ideas or designs without prior written permission is an infringement of the copyright protection of the work of each author.

© KCK Media Corp. 2022
Printed in the US

Directions for Hearables

There are many exciting areas of evolution currently in the audio industry, but no product category shows such levels of innovation, cutting-edge integration, and technology convergence as the true wireless stereo (TWS) earbuds category. For this issue, I tried to address that product class from a fundamental angle, which is voice integration. As I wrote this month, voice recognition and voice enhancement are defining features for hearables—currently the direction that is exciting everyone in the technology sector, particularly in the consumer electronics, mobile, and audiology markets.

We still don't know exactly all the applications and technologies that "hearables" will encompass, but we are starting to have some ideas. And they all look promising, starting with the fact that TWS earbuds are showing to be simultaneously the ideal audio front-end interface and also the ideal hub for biometrics and audio augmentation. Gradually, *audioXpress* has explored these topics in past articles, and we are far from exhausting the key areas in product development that are making TWS earbuds the fastest growing consumer product segment—already larger in volume than the smart speaker market, which was previously the fastest growing new product category.

Originally we intended to feature a separate article addressing the topic of voice and communications—as we did in our April 2021 edition. But the topic of TWS is showing so much activity on the voice front that we have condensed the coverage we had planned into a single Market Update. Basically, because the topics are so intertwined in current product designs—just look around and you will see that as people returned to the office, old-fashioned headsets used for office comms have now simply been replaced by those earbuds that everyone was already using at home and now love.

Voice and true wireless devices are closely related as any developer working in TWS earbuds and hearables can attest. In the article I wrote for this edition, I discuss the available hardware and software building blocks for these product designs, while also exploring the latest innovations in form and function, looking primarily at voice capture and recognition. We will still need to address topics such as active noise cancellation, RF modules, wireless streaming, and much more in future Market Updates.

And there's also the intense activity surrounding the new class of over-the-counter (OTC) hearing aids. While many consumer electronics companies are going to focus on hearing assistance and augmentation—features that don't necessarily meet the required qualifications as OTC hearing aids—the new earbuds will certainly perform much better than earlier concepts of personal sound amplification products (PSAPs).

And at the same time, we are also seeing all the audiology companies embracing the innovations being originated from this consumer segment. By integrating the best features from TWS earbuds, hearing aid companies can create more awareness within those with hearing loss, help remove social stigmas surrounding the use of hearing aids perceived as medical devices, and even foster new potential hybrid use-cases for compensating mild hearing loss, hearing safety, and hearing augmentation, directly applying to work requirements, or age-related risk factors.

And already on the horizon, there's the new field of augmented reality and immersive gaming, where true wireless stereo looks set to replace traditional headphones and headsets. With applications ranging from entertainment to corporate communications, there's a good chance that the added convenience—and the embedded power—of wireless earbuds will become an enabling component for the "metaverse" (whatever that will prove to be). Particularly as the ability to combine AI augmentation becomes a reality.

Product developers working in this field are certain to keep very busy for the foreseeable future.

J. Martins
Editor-in-Chief

The Team

Publisher:	KC Prescott	Associate Editor:	Shannon Becker
Controller:	Chuck Fellows	Graphics:	Grace Chen, Tori Zienka
Editor-in-Chief:	João Martins	Advertising Coordinator:	Nathaniel Black

Technical Editor: Jan Didden

Regular Contributors: Bruce Brown, Bill Christie, Bob Cordell, Joseph D'Appolito, Dimitri Danyuk, Vance Dickason, Jan Didden, Scott Dorsey, Gary Galo, Chuck Hansen, Richard Honeycutt, Charlie Hughes, Mike Klasco, David Logvin, Oliver Masciarotte, Nelson Pass, Christopher Paul, Bill Reeve, Michael Steffes, Nora Wong, Stuart Yaniger

OUR NETWORK

audioxpress | circuit cellar | LIS Loudspeaker Industry Sourcebook | Voice Coil

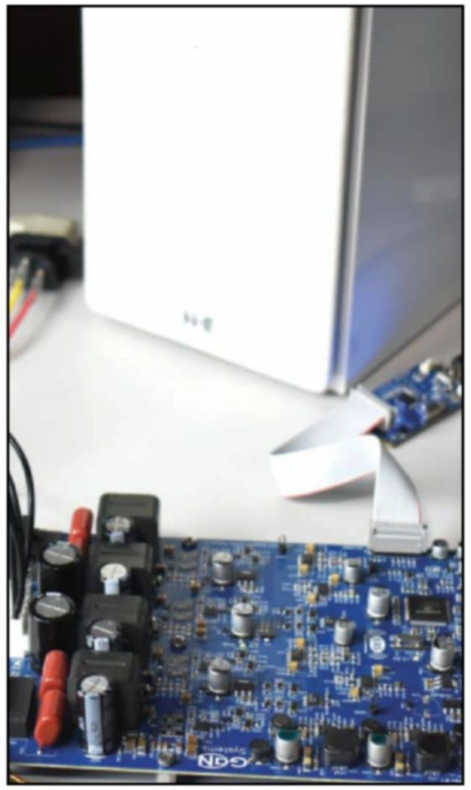
SUPPORTING COMPANIES

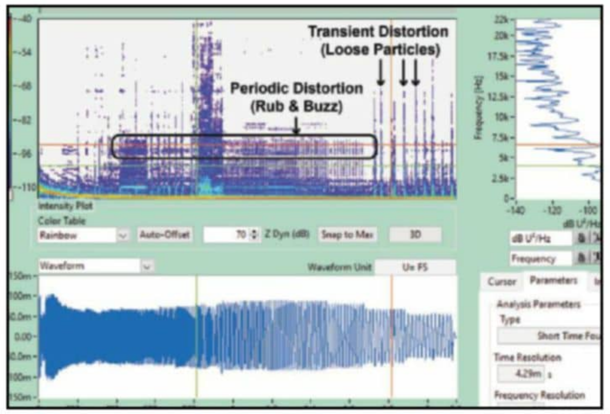
ACO Pacific, Inc.	41	KAB Electro-Acoustics	41
All Electronics Corp.	59	Knowles Corp.	13
Amplitrex Audio Products, LLC	55	Menlo Scientific, Ltd.	65
AXPONA	49	Parts Express International, Inc.	68
CEVA DSP, Ltd.	27	Profusion plc	45
COMSOL, Inc.	9	QuantAsylum USA, LLC	51
DSP Concepts	21	Redco Audio	59
Ellisys Corp.	23	SB Acoustics	35
GaN Systems	55	Syntiant	19
Hammond Manufacturing, Ltd.	3	Triad Magnetics	67
HEAD acoustics, Inc.	61	Western Electric	2
Hottinger Bruel & Kjaer, Inc.	37		

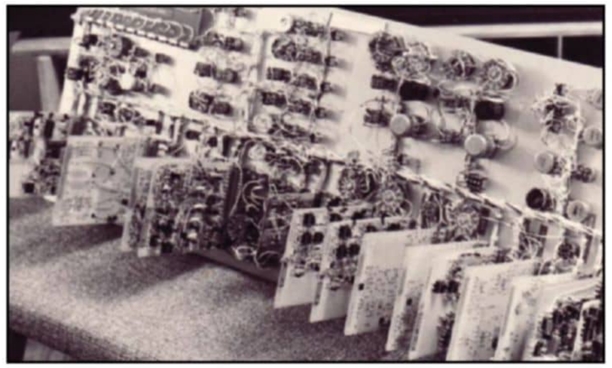
NOT A SUPPORTING COMPANY YET?
Contact Peter Wostrel (audioxpress@smmarketing.us, Phone: 978-281-7708, Fax: 978-281-7706) to reserve your own ad space in the next issue of the magazine.

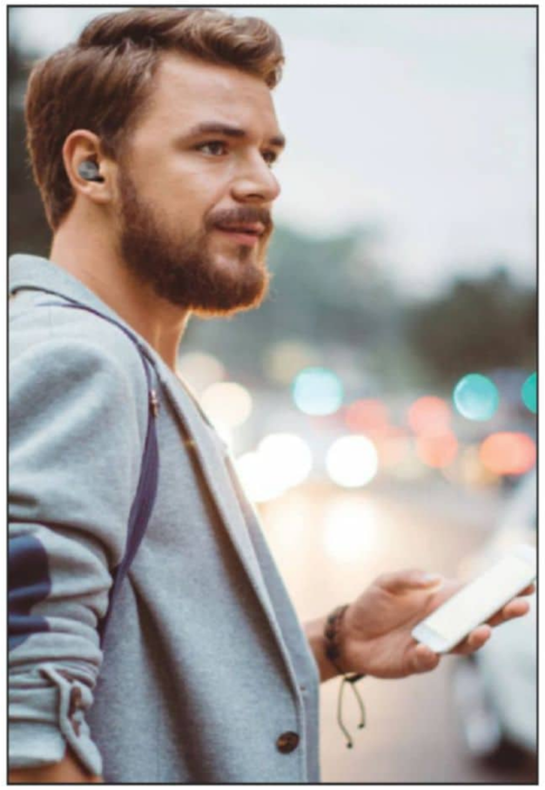
Contents

Volume 53 – No. 4 April 2022









Features

32 GaN Technology in Audio Power Amplification

By Stuart Yaniger

As he reviews a Class-D amplifier and switched-mode power supply (SMPS) evaluation kit from GaN Systems, Stuart Yaniger details the benefits for audio applications and shares his experiences with gallium nitride (GaN) power semiconductors.

42 The Evolution of Production Line Rub & Buzz Measurements

By Steve Temme

This article follows the development of production line Rub & Buzz detection, starting with the conventional methods that, although developed 25 years ago, are still valuable and well-used today, and ending with the latest perceptual algorithms.

52 An Introduction to Testing Vacuum Tubes (Part 1)

Reasons for Testing Tubes

By Mark Driedger

If you are a fan of tube audio, sooner or later you will need to test tubes. In this two-part article, Mark Driedger examines why and how we test tubes.

56 Building a Guitar-Controlled Synthesizer

The Sample & Hold Time Machine

By Ethan Winer

This article is the first chapter of a fascinating DIY journey that Ethan Winer takes through the world of synthesizers. More specifically, a guitar-controlled, all-analog synthesizer project that begins with a “time machine” sample and hold section.

Columns

SOUND CONTROL

28 Frequency Response of a Listening Room

By Richard Honeycutt

In this article, our audio expert argues that without standardization of the meaning of “frequency response” as applied to rooms, the term is probably not a useful metric from an engineering standpoint.

HOLLOW-STATE ELECTRONICS

62 Amplifier Comparisons Revisited

By Richard Honeycutt

After years of research and much debate, our columnist has determined that the superiority of tube or solid-state amplifiers is truly a matter of perception and is not necessarily a measurable or objective difference.

Market Update—True Wireless Stereo Earbuds

8 The Next Frontier in Voice Technology

By J. Martins

As any developer working in true wireless stereo (TWS) earbuds and hearables can attest, voice is a key feature in such designs. From voice communications to voice interfaces and voice enhancement, this article explores the available hardware and software building blocks for these product designs.

Edition Focus—Hearables in Focus

20 Putting the Smart in Smart Devices

The Future of Hearables Is Voice Controlled

By Probal Lala

Defined as ear-worn electronic devices, hearable technology is at the intersection of wearables and headphones and was designed for multi-purpose use, including voice transmission, sound amplification, and wireless communications.

24 The Acoustical Challenges of Designing True Wireless Earphones

By Andrew Bellavia

Consumer expectations for true wireless earphones are rapidly increasing, raising the bar on new designs. This article from Andrew Bellavia of Knowles Corp. explores the current trends as well as the acoustical challenges and solutions for supporting them.

Departments

4 From the Editor’s Desk

5 Client Index

Websites

www.audioxpress.com
www.voicecoilmagazine.com
www.cc-webshop.com
www.loudspeakerindustrysourcebook.com
www.circuitcellar.com
www.linuxgizmos.com

@audioxp_editor audioxpresscommunity
linkedin.com/company/audioxpress

Market Update—True Wireless Stereo Earbuds

By
J. Martins
(Editor-in-Chief)

The Next Frontier in Voice Technology

Any developer working in true wireless stereo (TWS) earbuds and hearables can attest that voice is a key feature in such designs. From voice communications to voice interfaces and voice enhancement, this article explores the available hardware and software building blocks for these product designs.

As both headphones and true wireless earbuds continue to gain augmented audio features, benefiting from the latest advancements in audio processing, active noise cancellation, spatial audio, and other innovations, we continue to see even more acceleration in building the basic blocks toward hearables. And voice—as we extensively detailed in our special report in the March 2022 issue of *audioXpress*—is a key feature in such designs. After all, voice recognition and voice enhancement are defining features for hearables.

Increasingly, TWS earbuds are praised for transmission call quality thanks to both the microphones positioned closer to the mouth and also the ability of these platforms to support beamforming and sophisticated digital signal processing (DSP) and source separation on-device. This contributes to making words become intelligibly delivered over a communication service, and for words to be recognized by speech processing systems and voice assistants.

These extremely compact and convenient systems are even proving to offer impressive results in very noisy environments, combining a tight integration of microphones and other sensors, together with DSP, even in very small form-factor earbuds. Add the fact that most premium models now feature audio transparency, enabling users to safely wear the earbuds on the move and in more day-to-day situations, interacting freely with other people, and it becomes clear why this product category is fast-becoming a primary voice front-end platform.

The latest Sony WF-1000XM4 true wireless earbuds continue to lead in the active noise cancelling category, showing an improved integration around its latest V1 Integrated Processor and Bluetooth 5.2 SoC, two noise-sensing microphones on the surface of each earbud, plus precise voice pick-up technology with an additional two microphones and a bone-conduction sensor for calls and voice assistants.

In many situations where manufacturers have considered adding voice interfaces and faced significant challenges due to distances, multiple users, and noise, TWS earbuds are proving to offer a very suitable solution in alternative to far-field voice capture. Adding the popularity of using earbuds for hands-free calls, this explains in part why there is so much focus with voice recognition for TWS earbuds—as a primary, personal interface.

The TWS Boom

Even the most basic and affordable TWS earbuds are meant to be used for much more than just listening to music. They combine basic headset functions, serving as a hands-free extension for smartphones during calls but also leveraging the bidirectional Bluetooth link (primarily) to enable voice control and dictation, to interactions with voice assistants. The fact that TWS designs combine both a microspeaker, one or more microphones, and a variety of sensors is a definition of such designs, and the key to its versatility, expanding quickly with on-device processing—and more recently with edge AI systems—into what we can define as the building blocks for hearables.

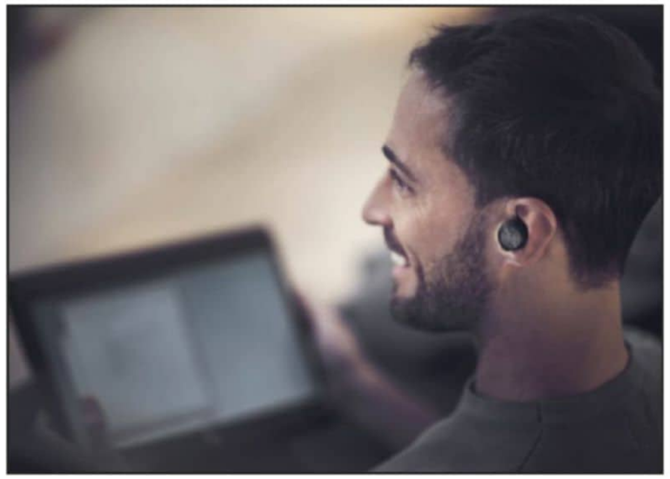
The functions that can be enabled by these sophisticated and fast-evolving in-ear devices are just starting to be explored and defined as new products with differentiating features become success stories—in this article we have selected a few purely consumer-level examples to illustrate these trends. No one questions the efficiency for voice capture of these devices any longer. This places true wireless earbuds as the next frontier in voice development—and currently the fastest growing product category for communication solutions.

Voice technologies have significantly evolved with smart speakers and the popularity of voice interface applications, particularly for

far-field use cases. But with TWS devices, the possibilities for applying sophisticated voice processing are simultaneously much more exciting and challenging—pushing research and development even further.

Currently, the volumes of TWS earbuds are encouraging those R&D investments, while many technology vendors are betting strongly on the creation of a new generation of cutting-edge computation devices for these low-power environments—with economies of scale that will eventually make it possible to spill over to other product categories. We've seen it happening with smartphones, then with smart speakers, and now with true wireless earbuds.

While many other features (e.g., active noise cancellation, spatial audio rendering, audio personalization and augmentation) are now



In 2021, Bang & Olufsen expanded its range of true wireless earbuds with the introduction of the Beoplay EQ with Adaptive Active Noise Cancellation, promoted as an effective solution for eliminating surrounding noise and allowing total immersion in music listening and calls. The new design uses an array of sensitive and highly directional, beamforming microphones, targeting working from anywhere.

In October 2021, Apple announced the third generation of AirPods, improving on its best-selling loose-fit, true wireless earbuds by adding many new updated features and deeply refining the design. While ANC remains exclusive to the more expensive AirPods Pro (in-ear) the new AirPods leverage the H1 chip to allow personalization and Adaptive EQ, improved microphones and acoustic mesh to help reduce the sound of wind, and adding AAC-ELD support—a superior speech codec that offers full HD voice quality—highly noticeable in natural communication for FaceTime calls and even for interaction with Siri.

being successfully added to earbuds in order to expand the market and find new use-cases, the reality is that after music reproduction, which remains the primary selling-point and use-case, voice capture and communication remain the second most essential features for earbuds.

The recent work-from-home and hybrid work requirements of the pandemic brought a focus on enhanced communications for this new emerging product class, to the point of the original white AirPods becoming almost an icon in web conference and streaming applications.

As stated in previous Market Updates, the Apple AirPods gained hundreds of millions of users forced to stay at home, including students. They also have been adopted by professionals and even TV reporters, which found the quality that could be achieved equal or superior to other headsets and devices available until now.

The reason for that is now well understood by all the hundreds of manufacturers that carefully dissected the different AirPods in order to copy the technologies, features, and even product design—sometimes just to find out that its performance depended upon an extraordinary level of integration of high-quality components and uniquely designed RF systems and antennas, plus proprietary cutting-edge system-on-chips (SoCs). While dedicated chips and more sophisticated ICs and integrated platforms are now generically available to any manufacturer, the next frontier then becomes the software integration, an area that most manufacturers don't even control, being limited to using readily available libraries and third-party code.

And that is currently the new frontier for this new product segment—finding the best platforms and designs to build original software integrations. Only Apple fully controls the complete equation, and that is why its business in this segment continues to grow even in face of exponentially increased competition.

As *audioXpress* also previously reported, it was expected that the market would quickly change in 2022 with the roll out of Bluetooth LE Audio, which was specifically optimized for TWS earbuds and hearing aids—but given the delay in the publication of the complete specifications, it seems that its impact on consumers is not likely to happen so soon. We will address Bluetooth LE Audio in a separate article, but we can already predict that in such a dynamic environment, we wouldn't be surprised to see companies proposing alternative proprietary technologies, or even looking at other standard efforts.

Sensor Fusion, Awareness, and Biometrics

Some key vendors in the TWS space are still focusing on the basics of audio, voice, and speech, allowing some evolution and differentiation to take place in those fundamentals. Among those, CEVA reports a clear interest from manufacturers and ODMs in solving challenges, such as barge-in for wake-word detection, and improved audio/voice processing based in context awareness. CEVA also offers solutions that enable earbuds to identify how they're being used, and in what type of environment, all while delivering a high-quality stereo experience and intelligently responding to multiple sensor streams with minimum latency.

In our previous Market Update about this topic, a year ago, we mentioned that CEVA had just announced Bluebud, a wireless audio platform aimed at standardizing DSP-enabled Bluetooth audio solutions for TWS earbuds, hearables, gaming headsets, and wearable devices. Bluebud was the first highly integrated solution announced to fully support Bluetooth 5.2 and LE Audio design, helping to support brands in the transition from Bluetooth Classic.

Running on the CEVA-BX1 DSP or an optional host CPU, Bluebud combines all the software and toolkits optimized for CEVA's DSP solutions, including audio codecs, voice assistants, spatial audio, in-ear detection, activity classification and full support for TensorFlow Lite Micro machine learning models. CEVA's optional software available for Bluebud also includes MotionEngine Hear for user control via tap and double tap, head tracking, in-ear detection and activity classification. For light AI inferencing workloads, Bluebud enables users to run TinyML models (e.g., a wake-word engine) to summon voice assistants or detect sound events for contextual awareness, using native support for the TensorFlow Lite Micro neural network framework.

In January 2022, CEVA announced Bluebud-HD, a pre-configured software package for its Bluebud wireless audio platform with the addition of immersive audio, wideband voice conversation, always-on voice assistant, assistive hearing, and IMU-based contextual awareness, all processed on the embedded CEVA-BX1 DSP.

CEVA also announced an upgrade to its SenslinQ software framework, upgraded for the Bluebud platform, to streamline the addition of third-party DSP algorithms and to facilitate the easy migration of existing software from other processor architectures.

The Bluebud-HD software package takes advantage of the on-board CEVA-BX1 DSP to consolidate a wide range of optimized HD audio, voice and context-aware modules, which are otherwise processed on separate DSPs or MCUs in non-CEVA architectures. This combination is a strong incentive for brands to standardize its developments around Bluebud to reduce the cost and the complexity of new wireless audio solutions.

In May 2021, CEVA announced that its RivieraWaves Bluetooth Dual Mode 5.2 Platform, licensed at that time by more than 15 wireless audio IC manufacturers and OEMs, had achieved Bluetooth SIG Qualification. In July 2021, shortly after the Bluetooth SIG announced the specification update, CEVA announced the general release of its RivieraWaves Bluetooth 5.3 complete hardware and software solution for new designs, and as a software-only upgrade for compatible existing designs, supporting all the latest features, including Isochronous Channels for LE Audio.

In January 2022, CEVA announced that its RivieraWaves Bluetooth Dual Mode 5.3 Platform had achieved Bluetooth SIG Qualification, along with its Generic Audio Framework (GAF) and LC3 codec.

Meanwhile, the Bluebud hardware platform has been upgraded to include CEVA's latest RivieraWaves Bluetooth 5.3 IP, and the CEVA-BX1 audio processor, together with all the required peripherals for wireless audio, in a compact low-power design.

Bluebud-HD also supports audio processing at 24-bit/96kHz, enabling high-resolution audio support. The fully integrated software package also includes CEVA's ClearVox, enhanced to support up to 48kHz sampling rate, for super wideband voice conversations; VisiSonics RS3D software for immersive 3D audio experiences; personalized hearing tuning for adaptation to each listener's hearing; and support for the LHDC high-definition audio codec by Savitech.

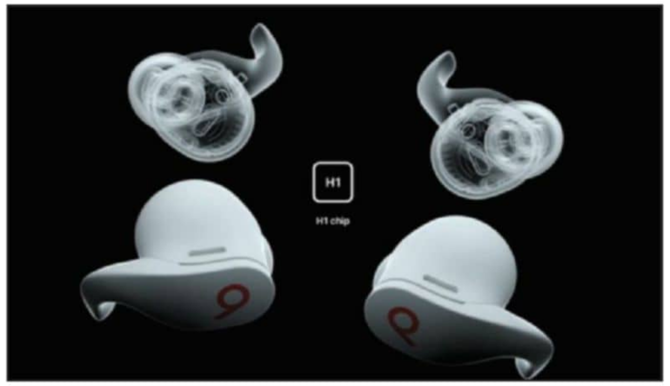
Launched in June 2021, the Beats Studio Buds are a powerful but affordable true wireless earbud design that delivers a high-quality audio experience. The new Beats Studio Buds add quality active noise cancellation, transparency, and improved voice capture in the tiniest, lightest, and more comfortable form-factor to date.

CEVA is also working with Bragi, the hearables pioneer from Munich, Germany, to expand options for manufacturers using the low-power CEVA-BX DSP architecture, in conjunction with Bragi OS, and its patented commercial app store. This enables technology partners and experts to offer software apps and services, such as voice activation, voice biometric, audio post processing and music services, licensed and downloaded directly to devices, and even enabling users to personalize their user experiences. CEVA also offers a range of its value-added software via the Bragi app store, including ClearVox for noise reduction, WhisPro for speech recognition, MotionEngine for sensor fusion, and the SenslinQ framework for contextual awareness.

By providing chip makers with mobile apps and a mobile SDK, Bragi is enabling audio brands to deliver new features, services, and user friendly apps with their products. This is a new approach to the TWS and hearables market and enables developers to introduce new audio features and services, as well as updates to those products already in the market, enabling new choices to the end user.

Capturing and Separating Voice

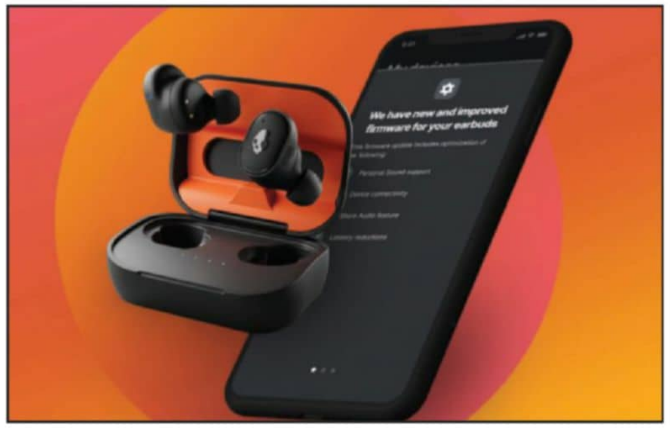
But the path toward hearables is also based on the concept that these in-ear devices will help support the integration of biometric



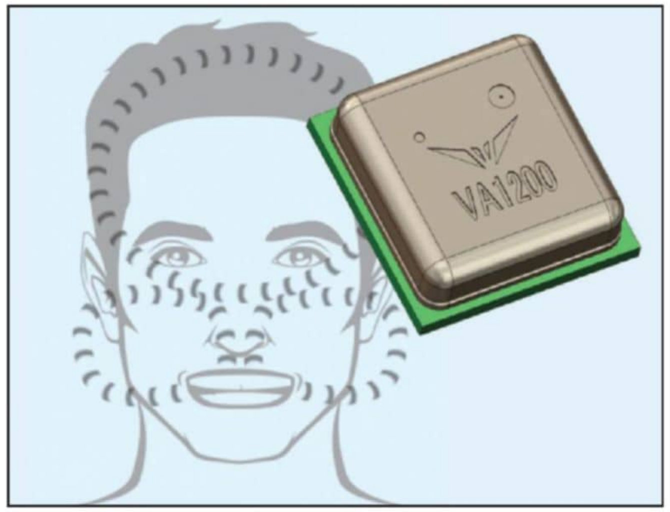
In November 2021, Apple surprised the industry with the introduction of the Beats Fit Pro advanced earbuds with ANC, retailing for just $199.99. The new custom-developed acoustic platform, engineered also for voice, features outward-facing and inward-facing microphones that work together to limit unwanted environmental sounds with targeted anti-noise, even as the environmental sounds change and adapt continuously, up to 200 times per second, to account for variances in fit and movement. Even without the stem of the AirPods, the Beats Fit Pro can compete in call quality, leveraging beamforming microphones that target the user's voice, acting in coordination with the internal mic and voice accelerometer to filter out ambient noise and create a natural conversation experience.

monitoring, converting TWS earbuds in effective personal health devices, complementing other wearables. And there's obviously also the impact that these devices are having in hearing health and audiology sectors, with TWS earbuds effectively being converted into hearing assistance devices, while hearing aids are increasingly emulating the features, technologies, and even the form factors from the consumer market.

In all these applications, voice front ends and voice processing remain at the core and the industry is looking for more effective components and technologies. As an example, in 2021 piezoelectric MEMS microphone pioneer Vesper introduced the VA1200, which the company calls the world's first analog piezoelectric voice accelerometer (a.k.a. bone conduction sensor). This low-power vibration sensor can be used in true wireless earbud designs to pick up the wearer's own voice through bone conduction. Using the VA1200 in conjunction with a standard microphone and a sensor fusion solution, the application can achieve superior background and wind noise reduction, according to Vesper.



Skullcandy's voice-driven platform, Skull-iQ Smart Feature Technology is powered by BragiOS, he was perfected in a development effort between Skullcandy and Bragi. The first products to use the integration—the new Grind Fuel and Push Active TWS designs—support intuitive voice commands directly to a device's personal voice assistant without any additional setup.



Vesper introduced the VA1200 analog piezoelectric voice accelerometer, which can be used in true wireless earbud designs to pick up the wearer's own voice through bone conduction. Using the VA1200 in conjunction with a standard microphone and a sensor fusion solution, the application can achieve superior background and wind noise reduction.

The VA1200 is the first non-microphone product based on Vesper's patented transducer technology and validates the notion that the company's piezoelectric MEMS platform can be used for a broad array of sensor types.

The piezoelectric voice accelerometer picks up the vibration of the user's voice through the skull. By design, it filters out background noises (e.g., wind noise, music, subway noise, and crowd noise). Of course, the fidelity of the user's voice that is captured through bone conduction is not as faithful as the signal captured by a traditional microphone, but it can be used to improve the filtering that is required to separate the voice from any background noise, even in challenging noisy environments. The VA1200 is already being integrated in new TWS designs where the VA1200 sensor also enables advanced features, including high-fidelity voice pick-up, Automute, Tap Detection, and Voice Authentication.

As *audioXpress* addressed in our Market Update on Voice Capture and Voice Interface Systems published in the March 2022 issue, microphones are a key component to enable the evolution of voice processing and communications in TWS earbuds. As the microphone's specifications improve, their effectiveness is also expanding thanks to designs using a larger number of microphones and enabling new beamforming techniques in combination with DSP, which are key to the direct integration of voice assistants at the edge.

As market research firm Yole Développement details in its 2021 report on Microphones, Microspeakers, and Audio Processing, "The audio industry is entering a new phase, with new technologies, more software and more algorithms enabling improved performance and increased value. Microphones are increasingly being integrated into all types of consumer devices to support smart assistant integration and immersive audio experiences.

Integration of audio processing and AI computing functions inside APUs or SoCs for specific end-systems is another major trend. In addition, TWS earbuds are becoming a commodity product, boosting sales of audio-related components."

As Alexis Debray, Senior Analyst at Yole Développement writes in the report: "Audio is the new must-have interface to interact with phones, speakers, cars, and even homes. Audio is also bringing new immersive experiences with ANC, and 3D audio, but also sound transparency to better connect with reality."

According to Yole, microelectromechanical systems (MEMS) microphones are the components forecast to have the second-highest growth rate between 2020 and 2026, a 10.5% CAGR from $1.4 billion USD to $2.5 billion USD, as microphones become an essential component in most end-systems, enabling a multitude of functionalities and enhancing the audio experience.

As a reminder, we previously discussed the state of the art in MEMS microphones our February and March 2022 issues, and the signal processing that takes place on the voice that is captured in our DSP Market Update in the November 2021 issue.

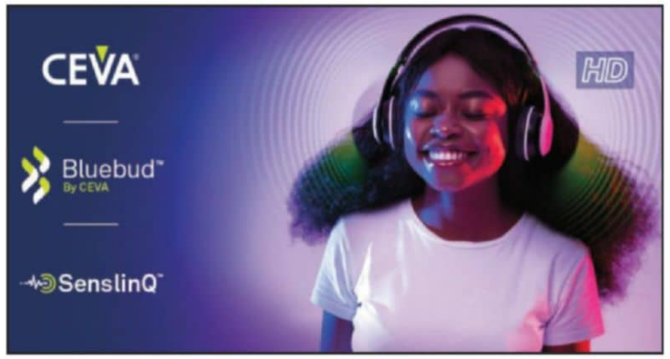
Chips and DSP for TWS

All the key software and algorithm companies have been extremely busy over the past 12 months to expand their market reach by supporting the most important platforms being used in TWS earbuds. We have mentioned CEVA already, and we need to mention Qualcomm, which has been aggressively expanding its solutions for this thriving market.

With its previous generation of QCC51*xx* and QCC30*xx* series SoCs for Bluetooth true wireless earbuds development, Qualcomm enabled virtually every single mobile and CE brand to create products able to compete with Apple's massively successful AirPods. Then Apple introduced its AirPods Pro with ANC and its market share expanded even more.

Qualcomm responded by launching a new generation of SoCs combining hybrid ANC and other technologies to enable brands

Technics' EAH-AZ60 and EAH-AZ40 true wireless earbuds are equipped with JustMyVoice voice detection technology, which provides improved clear voice communication during calls. The EAH-AZ60 model uses a total of eight mics on the left and right to power Dual Hybrid Noise Cancelling Technology, combining feedforward (FF) and feedback (FB) methods, as well as digital and analog processing for best-in-class ANC.



At CES 2022, CEVA announced that it has enhanced the software offerings and tool kits for its highly successful Bluebud wireless audio platform with the addition of Bluebud-HD, a pre-configured software package for the Bluebud platform, featuring immersive audio, wideband voice conversation, always-on voice assistant, assistive hearing, and IMU-based contextual awareness, all processed on the embedded CEVA-BX1 DSP.

to design new competitive TWS products. This next generation of ultra-low-power Bluetooth SoCs were designed to enable improved robustness and connectivity, increased battery life, integrated dedicated hardware for Hybrid Qualcomm Active Noise Cancellation (Hybrid ANC), Voice Assistant support, and premium wireless sound and voice quality.

Next up, in 2021 Qualcomm announced Snapdragon Sound, introducing a new end-to-end approach to promote high-resolution digital audio and enhanced voice communications across devices. Snapdragon Sound combines multiple Qualcomm technologies, including all its Bluetooth Audio QCC SoC families (QCC514*x*, QCC515*x*, and QCC3056), Aqstic Audio Codec and Aqstic Smart Speaker Amp, the latest aptX Adaptive audio codecs, TrueWireless Mirroring, ANC, aptX Voice (32kHz super wideband voice quality), and new 3D/spatial audio tools. Everything needed for a manufacturer dependent on Google's Android OS to deliver solutions that seem to compete head to head with what Apple offers—differentiated with a more complete set of features and options for developers.

For OEMs, Snapdragon Sound means access to pre-certified solutions in both devices and accessories categories, to reduce development time and shorten time-to-market. For audio brands, the system enables product differentiation, boosting the Android ecosystem as a more vertically integrated experience. And finally, for end-users adopting Google's ecosystem, a perception of choice with verified quality and performance—something that never existed in the Android environment.

To this all-encompassing strategy, Qualcomm's Snapdragon Sound (a direct reference to its family of processors) promises to be an end-to-end systems approach around audio. Also important to mention that, shortly following this announcement, Qualcomm introduced a CD-quality lossless option over Bluetooth, expanding aptX Adaptive with the aptX Lossless codec for Bluetooth wireless audio, combining wireless link and RF optimizations. This strategy makes sense given that Qualcomm confirmed a partnership with Amazon, and its new Amazon Music HD service, which includes lossless streaming.

To facilitate development of new designs based on Qualcomm Snapdragon Sound, Goertek launched a new earbud reference design powered by the Qualcomm QCC5141 Bluetooth Audio SoC. The Goertek Rox reference design is optimized for use in TWS earbuds and hearables and is designed to deliver high-resolution music, improved voice call quality, and lag-free gaming—with superior connectivity and extended battery life. Each earbud of the Rox reference design uses three Goertek high-SNR microphones and one Super Balance Speaker (SBS) driver to support high-performance ANC.

Goertek also supports always on wake-word detection for Amazon Alexa or Google Voice Assistant, and offers ANC and EQ tuning to help OEMs reduce development time and cost.

Still on the topic of enabling processing platforms, with its Apollo3 and Apollo4 family of ultra-low-power microcontrollers and SoCs powering many voice-enabled connected devices, Ambiq is another company that now offers a solution for manufacturers looking to introduce voice-command in TWS earbuds. The new Ambiq Voice-on-SPOT (VoS) Kit offers a complete, always-on, always-listening, voice-to-cloud solution for local voice commands and/or voice assistant integration at both the MCU and the system levels. The VoS Kit integrates Ambiq's hardware and software with peripherals and third-party IP, including signal processing using DSP Concepts' Audio Weaver, Sensory VoiceHub, and Retune DSP VoiceSpot (now part of NXP, see more about that in this article). While its targeted applications typically include remote controls, gaming controllers, smart watches, and health/fitness trackers, Ambiq is increasingly expanding to TWS earbuds and hearables, supporting advanced voice assistant, and even real-time translation implementations.

A strong contender in this field is Cadence Design Systems, whose Tensilica HiFi DSP configurable processors are present in most of today's audio, voice, and speech processing applications. Recently, the company announced the new Cadence Tensilica HiFi 1 DSP, optimized for ultra-low-energy implementations. Delivering increased voice- and music-processing performance with optimal neural network capability in a compact footprint with ultra-low energy, the Tensilica HiFi 1 DSP is now the smallest and lowest power member of the HiFi DSP family.

As required for TWS earbuds and hearing aids, these processors accelerate AI function and machine learning workloads for voice applications, along with running codecs and DSP/AI algorithms for pre-processing and post-processing. These processors were also the first solution to fully support the latest Low Complexity Communications Codec (LC3 and LC3plus) licensed from Fraunhofer.

The HiFi 1 DSP now delivers ultra-low energy encoding and playback of LC3 and other Bluetooth codecs and ultra-low energy keyword spotting for voice wake-up, in the smallest footprint HiFi DSP. Compared to the Cadence HiFi 3 DSP, one of the most popular audio DSPs in the industry for these target applications, the HiFi 1 DSP offers 11% to 16% lower area, 60% to 73% greater cycle and energy efficiency for ML-based wake-word or keyword spotting and person detect applications. It also offers greater than 18% cycle efficiency and 14% energy efficiency for LC3 decoding.

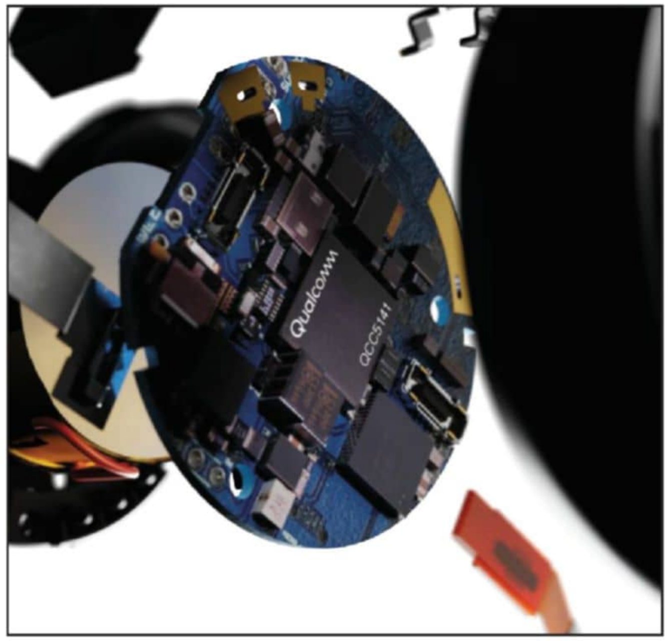
Voice Processing Technology for TWS

Available for many other platforms, developers relying on Qualcomm's audio SoCs are also able to offer improved voice clarity in noisy or windy environments using Alango's OnlyVoice advanced technology for true wireless earbuds and high-performance headsets. Designers adopting the Qualcomm QCC5141 can now offer this advanced voice acquisition solution relying on sensor fusion technology for devices incorporating acoustic microphones and a vibration sensor.

Part of the multiple digital signal processing solutions developed by Alango Technologies, OnlyVoice improves communication intelligibility in noisy environments by employing an in-ear signal source, which due to its positioning is isolated from ambient noise. The combination of the in-ear sensor and external beamforming array captures clear, full-spectrum speech surpassing the performance limitations of modern-day headphones. OnlyVoice also includes Voice Activity Detection (VAD) functionality allowing low-power, always-listening operation of voice-activated devices.

Alango has been a long-time member of the Qualcomm Extension Program, successfully providing Alango DSP algorithms running on Qualcomm Technologies' Bluetooth SoCs. Already available for the Qualcomm QCC5141 platform is Alango VoiceFirst (previously known as SEN or Sound Effect Normalization), a unique solution developed to improve speech intelligibility in movies, television shows, and sports broadcasts through de-emphasis of background sounds and music. Besides increasing dialogue intelligibility, VoiceFirst can also reduce the user's fatigue from exaggerated sound effects, making content more suitable for older adults.

Grell Audio, the company founded by German headphone engineer Axel Grell, announced its first product, the Grell TWS/1 earbuds, implemented over a Qualcomm Bluetooth 5.2 platform and offering a frequency-selective ANC function, filtered by a patent-pending Noise Annoyance Reduction (NAR) psycho-acoustic model, bundled with Sonarworks' SoundID for calibration and personalization.



Qualcomm's leading Bluetooth Audio QCC SoC families (QCC514x, QCC515x and QCC3056) currently power a large number of true wireless earbuds from leading brands such as Bang & Olufsen, Belkin, Cambridge Audio, Cleer, Earin, Edifier, Jabra, JVC, Sennheiser, Yamaha, and many others.

The Snapdragon Sound technologies and components will address a variety of all application scenarios but will certainly enable new-generation TWS earbuds with high levels of performance and close integration. The Snapdragon Sound systems program will be tested for interoperability and certification directly by Qualcomm Technologies in a dedicated test facility in Taiwan.

Alango also offers its Extended Voice Communication Package (eVCP), a universal software suite of digital signal processing technologies for voice applications, running on DSP Group's processors (now Synaptics), as well as platforms from Cadence, CEVA, NXP, Knowles (AISonic), and others. And manufacturers can fully explore the seamless integration of OnlyVoice and other Alango VCP technologies, which include noise suppression, acoustic echo cancellation, dynamic range compressor, automatic volume and equalization control, speech enhancement, and several other functions.

For brands interested in combining voice communications, voice interfaces, and hearing-assist features in either consumer hearables or OTC hearing aids, Alango recently announced that it's now offering licensing options of its Hearing Enhancement Package (HEP) software technologies to OEMs and ODMs. These software technologies, which Alango perfected for its own range of affordable BeHear designs available direct-to-consumer, allow full-featured customer-fit hearing personalization and amplification, enabling

Goertek's earbud reference design helps manufacturers looking to implement Qualcomm Snapdragon Sound technology speed their time to market. Powered by the Qualcomm QCC5141 Bluetooth Audio SoC, the Rox reference design is optimized for use in TWS earbuds and hearables and to deliver high-resolution music, crystal-clear voice calls, and lag-free gaming—with superior connectivity and extended battery life.

the rapid development of advanced hearables in new designs and form-factors. Additional technologies for sound personalization can also be integrated (e.g., situational awareness), to pass important sounds through while listening to audio, and the ability to adjust the volume and spectrum of the audio depending on ambient noise conditions, automatically.

Voice Software Integrations

On a similar approach of wide-reaching initiatives, DSP Concepts, the creators of Audio Weaver, the audio platform powering many millions of consumer products, continues to extend support to many new technologies that aim to improve voice capture and recognition, hearing personalization and assistance. Recent announcements on this front include collaborations with Sensory, Fluent.ai, and Mimi Hearing Technologies. Mimi offers its app-based hearing test and compensation technology, and allows audio manufacturers to provide a personalized listening experience on TWS earbuds.

Mimi also employs a smartphone app and digital signal processing technology built into true wireless earbuds to administer a user-conducted hearing test that optimizes the audio based on the individual's hearing and yields instant results. Mimi's technology gauges the user's ability to distinguish sounds in the presence of interfering noise and develops a personalized restoration profile, called a Hearing ID, to allow personalized hearing correction across multiple devices. beyerdynamic, Skullcandy, Cleer Audio, Philips, Loewe, and more brands have already integrated Mimi's technology into millions of units to date.

Another interesting development to enable advanced designs on this front is the Audio Weaver integration with CloudNav, a California company specializing in sensor fusion. Originally focused on solving multiple known algorithmic problems in audio, motion tracking, and location, CloudNav offers a suite of software products that help support advanced integrations to feed context- and location-based awareness and activity detection in low-power designs (e.g., wireless earbuds). Its TinyFusion solution is the world's lowest power and smallest memory footprint 9-axis sensor fusion (gyroscope, accelerometer, and magnetometer) for challenging environments and is both sensor and processor independent.

And since we mentioned Retune DSP, the Danish provider of DSP software and algorithms for voice control and communication applications was one of the almost invisible acquisition stories in 2021, allowing NXP Semiconductors to add three powerful voice DSP products.

Retune DSP has built a strong reputation as a one-stop shop and full-service software supplier, delivering a complete voice control stack, including multi-microphone beamforming, multichannel acoustic echo cancellation, and wake-word engines. Retune DSP's solutions provide highly accurate, small-footprint, low-power voice processing.

When developers required extremely high levels of detection accuracy within tiny model sizes, Retune DSP stepped up to deliver the right voice solution for each application. The three basic Retune software tools include Conversa, a high-performance, multi-microphone, full-duplex voice processing software suite; VoiceSeeker a state-of-the-art multi-microphone voice control audio front-end DSP solution; and VoiceSpot, a highly accurate wake-word engine using machine learning. Combined with NXP's low-power platforms and development tools, these Retune tools can be a valuable option for voice implementation in TWS earbuds.

Also targeting the low-power requirements of true wireless earbuds, Belgium company 7 Sensing Software is working with Analog Devices to incorporate its advanced embedded audio DSP and machine learning solutions. 7 Sensing already has significant experience working with manufacturers on advanced voice processing and hearing augmentation features for hearables. In those implementations, 7 Sensing is able to support multiple platform cores from Qualcomm, Cadence, and ARM. The company's audio DSP and AI-optimized libraries are key embedded software technologies, helping to solve the challenges of voice quality calling with wireless earbuds, especially in noisy environments. 7 Sensing's Pure Voice software tackles that challenge by leveraging bone vibration sensors and in-ear microphones. Pure Voice intelligently blends the signals from multiple sources for crystal clear conversations, with speech transmitted cleaned from background noise.

The collaboration with Analog Devices will perfectly complement the company's latest generation of low-power codecs, such as the ADAU1788, the ADAU1850, or the more powerful ADAU1860 with three inputs and two DSPs for more demanding ANC earbud designs.

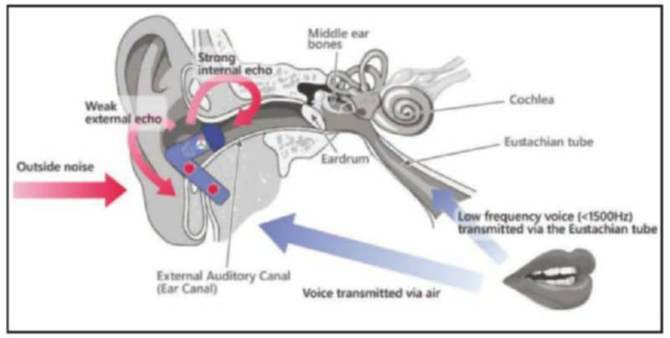
Voice and AI

While we don't intend to expand this Market Update into the field of artificial intelligence (AI) and machine learning (ML) tools, which we intend to explore further in a future dedicated report, it's important to mention that voice applications are increasingly based on these advanced edge AI platforms.

And leading the charge for very low power requirements, such as TWS earbuds, is Syntiant. Shortly after announcing its latest and most powerful Neural Decision Processor NDP102, in September 2021, Syntiant unveiled its TinyML Development Board, an easy-to-use developer kit aimed at both technical and non-technical users for building ML-powered speech commands, wake-word detection, acoustic event detection and other sensor use cases. The new machine learning development board enables developers to build low-power voice, audio and sensor applications using Edge Impulse's Embedded ML platform, and is a small, self-contained system, with onboard digital microphone and 6-axis motion sensor.

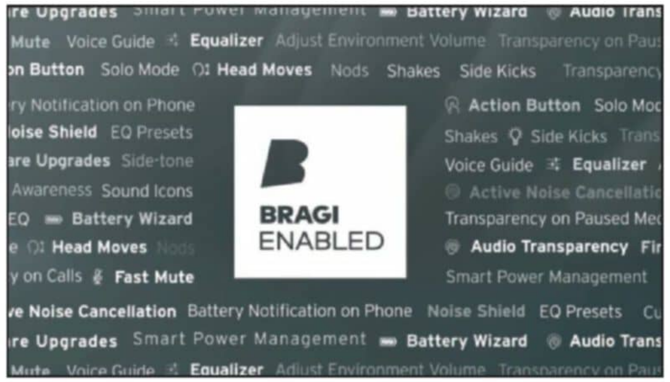
Already in 2022, Syntiant announced a new reference design that will enable ODMs and OEMs to easily deploy these ultra-low-power, edge AI voice processing applications in true wireless earbuds. The solution enables multilingual custom wake-words and speech commands, supported by data training and modeling, providing a complete, end-to-end solution for edge AI processing. Among its first ODM clients was Wondersound, a Shenzhen, China-based company that has introduced hearables for leading consumer brands, including Anker, Baidu, boAt, Noise, and OnePlus.

It shouldn't be a surprise that Chinese companies are leading the charge in advanced AI-based language translation and voice assistant solutions. China offers an ideal market for a test-bed of these concepts, with a thriving environment of hardware and software technologies that are more evolved than anywhere else. When it comes to supporting Chinese language voice recognition,



Alango OnlyVoice improves communication intelligibility in noisy environments by employing an in-ear signal source, which due to its positioning, is isolated from ambient noise. The user's speech is acquired with either an in-ear microphone via the eustachian tube or an in-ear accelerometer via bone conduction. The in-ear signal is combined with the external microphone's beamforming signal and intelligently mixed, since the frequency limit of internal voice pickup is effectively 1.5kHz. OnlyVoice also includes Voice Activity Detection (VAD) functionality allowing low-power, always-listening operation of voice-activated devices.

7 Sensing Software offers its Hearing Augmentation and Pure Voice processing solutions for wireless earbuds and headsets running on ultra-low-power platforms from Qualcomm, Cadence, ARM, and now Analog Devices.



In 2017, Bragi announced its Bragi OS for hearables supporting multiple service integrations, such as Amazon Alexa, and embedded AI to automatically detect activities and context. In 2021, the first Bragi Enabled products reached the market from Skullcandy and Klipsch. Among the many features currently enabled by Bragi OS, on-device Voice Control is one of the main focus.

things are intrinsically more complex than the English language, with a complex weave of variants and dialects, while consumers require (and rely upon) multilingual translation and transcription more than in any other market.

A leading company in that space is Voice AI, offering identity authentication with its Voiceprint recognition system, based on deep learning convolutional neural network. Voice AI specializes in these complex voice implementations that are able to recognize multiple users, expanding speech recognition to speaker (talker) recognition, pronunciation, content, and meaning. This type of technology is vital when implementing multilingual translation in a natural conversation.

Klipsch's latest T5 II True Wireless ANC earbuds are the result of an extended collaboration with Dirac for digital audio processing, and Bragi's software suite with embedded AI. The design takes advantage of a total of six beamforming microphones to power the ANC/Transparency features and calls. Among other features, the built-in Bragi operating system allows for hands-free, advanced gesture-control, Noise Shield, to protect the user from unexpected loud noises, and Voice Awareness, allowing users to hear their own voice naturally while on calls.

Also working at the intersection of AI and voice, DSP Concepts is expanding its initiatives to support advanced edge AI implementations directly in Audio Weaver, working in partnership with companies such as Sensory and Fluent.ai.

Understanding that voice-enabled products must be able to adapt to different languages and regional accents, and that developers struggle with the time and costs to deal with the challenges of voice integration, DSP Concepts provides the solutions for the basic problematic acoustic environments, latency and accuracy of voice recognition, interfacing directly with existing processors and audio front-ends. The next level of requirements needs to be addressed by voice products that allow deployment across multiple regions and support trained models and AI/ML.

To meet these demands, DSP Concepts provides the Audio Weaver platform and the TalkTo audio front end, combined with the edge-based Fluent.ai Air automatic intent recognition engine.

TalkTo combines advanced signal processing techniques to deliver clean audio signals to voice assistants and speech recognition

In 2021, Amazon introduced updated Echo Buds, showing how the company that pioneered voice assistants was able to improve access to its own Alexa voice assistant in true wireless earbuds. Achieving a 20% smaller and lighter design, Amazon introduced improved active noise cancellation and improved voice control.

Sources

7 Sensing | www.7sensingsoftware.com

AIStorm | www.aistorm.ai

Alango Technologies, Ltd. | www.alango.com

Ambiq Micro, Inc. | www.ambiq.com

Analog Devices, Inc. | www.analog.com

Bragi | www.bragi.com

Cadence Design Systems, Inc. | www.cadence.com

Ceva, Inc. | www.ceva-dsp.com

CloudNav, Inc. | www.cloudnav.net

DSP Concepts | www.dspconcepts.com

Edge Impulse | www.edgeimpulse.com

Goertek | www.goertek.com

Knowles Electronics, LLC | www.knowles.com

NXP Semiconductors | www.nxp.com/voice

Qualcomm | www.qualcomm.com

Retune DSP | www.retune-dsp.com

SAR Insight | www.sarinsight.com

Sensory, Inc. | www.sensory.com

Skullcandy | www.skullcandy.com

Syntiant Corp. | www.syntiant.com

Vesper Technologies, Inc. | www.vespermems.com

VoiceAI Technologies Co., Ltd. | www.voiceaitech.com

Yole Développement | www.yole.fr

engines, and Fluent.ai Air offers edge-based universal language and accent support, detecting intent without the need to connect to the cloud and without converting speech to text.

Like TalkTo, Fluent.ai Air is a scalable technology and predefined commands can be triggered with variable or synonymous phrasing, using a "direct intent" model that can be supported by low-power devices utilizing a smaller vocabulary and less variable command phrasing.

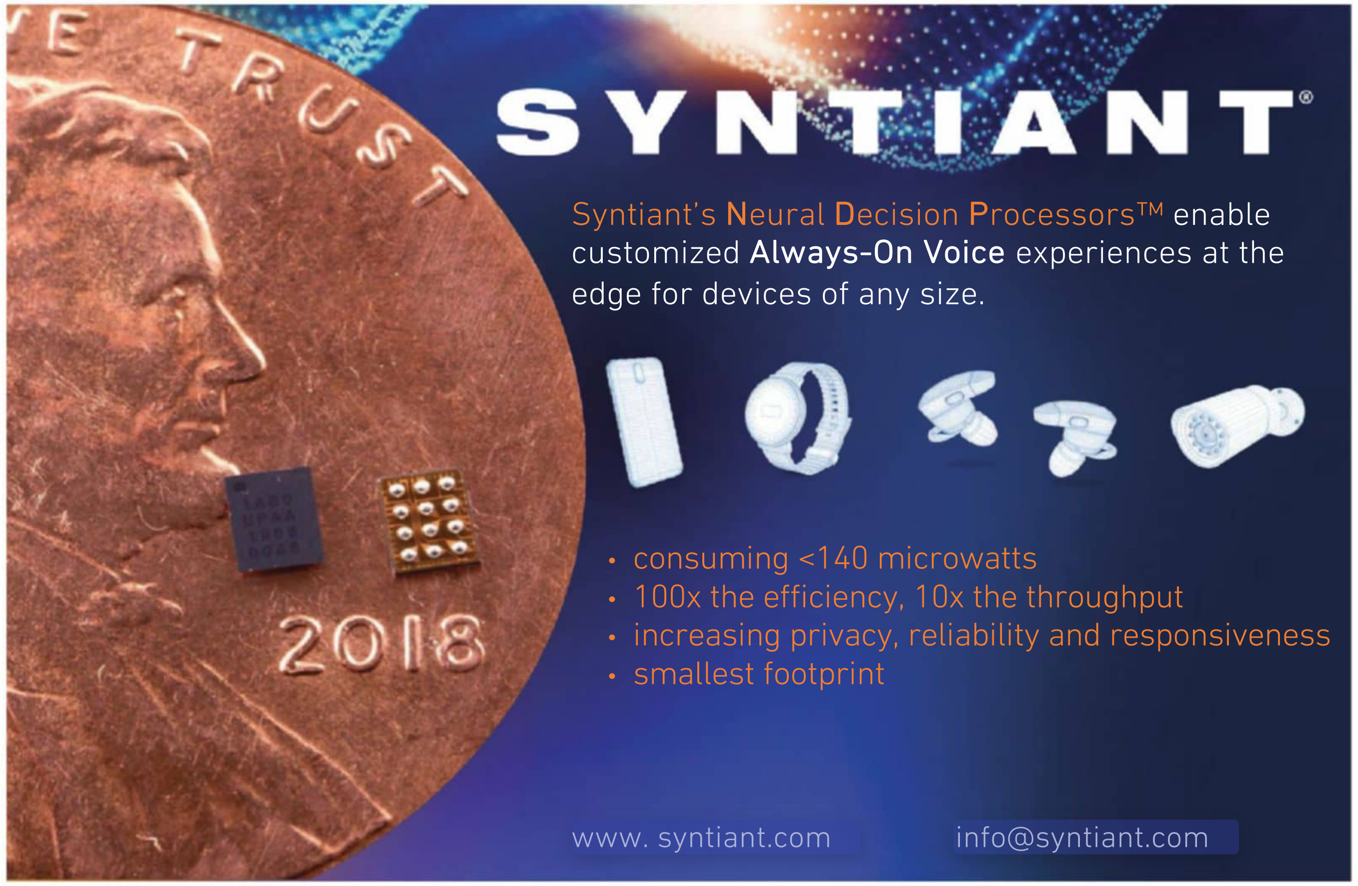
Technologies, such as Fluent.ai Air, point forward to a future where AI devices will be intelligent in and of themselves and will no longer need to depend on an Internet connection.

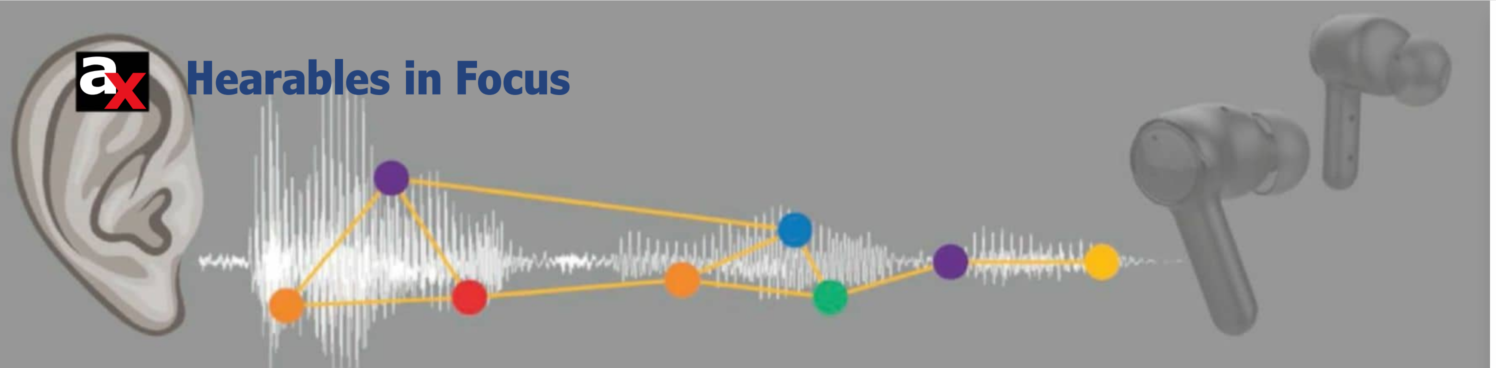
DSP Concepts is also working with Sensory to offer integrated solutions for wake-word detection, command and phrase spotting, natural language processing, and sound identification using embedded machine learning. Sensory's machine learning solutions—TrulyHandsfree for wake-word detection and phrase spotting, and TrulyNatural for continuous speech recognition—represent a voice user interface that is private, fast, and requires no cloud computing.

We will mention a final example of platforms for edge AI audio solutions recently announced through the cooperation between Knowles—with its AISonic processors—and AIStorm, a pioneer of AI-in-Sensor SoCs for edge devices. Combined with Knowles' audio signal processing technology and advanced micro-acoustic microphones and speakers, AIStorm's analog and mixed-signal audio data can be directly processed in low-power TWS earbuds.

Syntiant, the company focused on making edge AI a reality for always-on applications, announced a new reference design that will enable ODMs and OEMs to easily and quickly deploy ultra-low-power, edge AI voice processing applications in true wireless earbuds.

These small form-factor devices are challenged to perform advanced signal processing tasks, such as noise cancellation, beam forming, keyword spotting (KWS), speaker identification, sound classification, and other algorithmic sound enhancements at very low power levels. The combined AIStorm and Knowles solution, which can also be integrated with Fluent.ai's embedded speech recognition, promises to deliver an advanced approach for these complex design requirements, delivering on the vision of what we can characterize as smart hearables.





Putting the Smart in Smart Devices

The Future of Hearables Is Voice Controlled

Defined as ear-worn electronic devices, hearable technology is at the intersection of wearables and headphones, and was designed for multi-purpose use, including voice transmission, sound amplification, and wireless communications.

By
Probal Lala
(Fluent.ai)

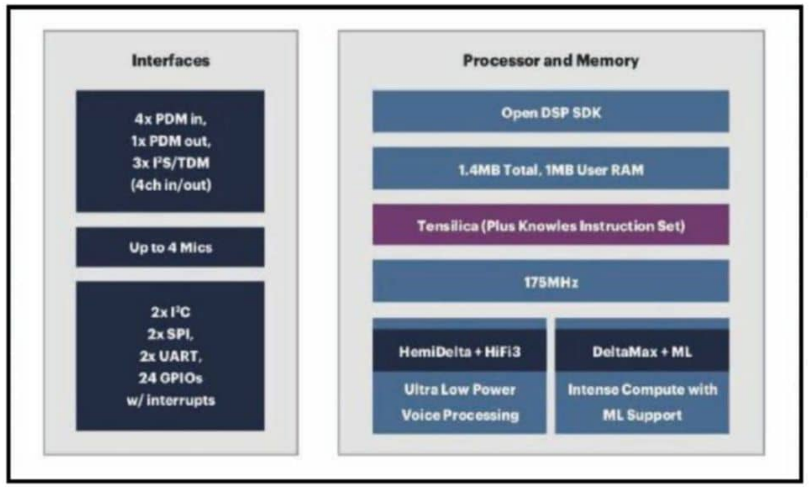
Increasingly used for voice communications on the move, consumers expect hearable devices to work in noisy public environments (e.g., their workplace, shopping malls, and on public transportation) in addition to remote and home environments. As a result, hearables are expected to provide the portability and convenience that consumers demand from their electronic and smart devices; however, the technology faces limitations. Most voice-enabled hearables on the market require a smartphone app or Wi-Fi connectivity to operate and existing offline command sets are relatively small with global adoption hampered by accent or language limitations of the technology.

The work-from-home phenomenon, easing pandemic, and growing consumption of mobile games and streaming with the adoption of 5G is expediting growth of the market, and the evolution of true wireless stereo (TWS) technology is leading the charge in advancements to meet these consumer demands. With this in mind, every audio and technology company will be paying serious attention to the hearables market in the years to come.

TWS is a term used to describe any technology that offers a way to listen to high-quality stereophonic sound wirelessly and Bluetooth technology is typically used for wireless transmission of sound in TWS devices. TWS earphones have developed rapidly since their introduction in 2015 and took off in mainstream consumer markets when Apple introduced the AirPods in 2016.

The global hearables market size was valued at $21.20 billion in 2018, and Allied Market Research projects it to reach $93.90 billion by 2026, growing at a compound annual growth rate (CAGR) of 17.2% from 2019 to 2026. According to Strategy Analytics, overall adoption of TWS headsets alone will grow 54% year-over-year in 2022.

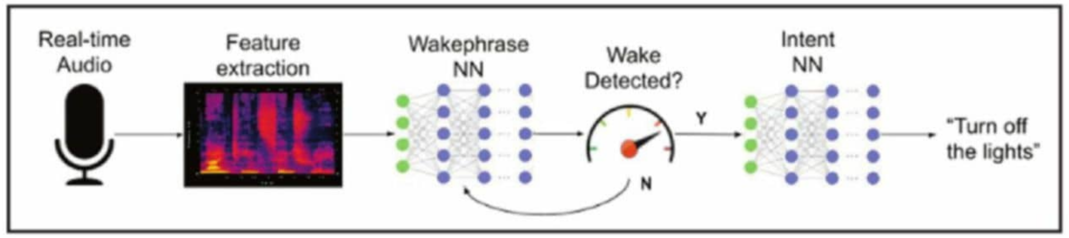
The potential of TWS has not yet been fully realized and innovation in several key technology areas, including Edge AI, ensures that the growth and evolution of these devices will continue.



Features of Knowles AISonic Audio Edge Processor IA8201: Dual Core

Limitations of Current Hearables on the Market

The term "hearables" was first coined by technology analyst Nick Nunn in a 2014 blog post in which he declared the ear as "the new wrist" for wearable technology. At the time, in-ear devices capable of interacting with smartphones or other technology had yet to dominate the market. Hearable devices now have the ability to provide enhanced listening experiences; however, the future of hearable technology is just beginning to emerge.



Fluent.ai Wakeword + Air speech-to-intent engine

Advancements in hearable technology are challenged by the need for small footprint software. Hearables need to integrate a speaker, a microphone, and sensors, meaning only small audio integrated circuits (ICs) will fit in the device. Despite massive improvements in microchip and sensor technology, embedding various technologies into smaller earbuds remains challenging, and the performance of cloud-dependent hearable devices is limited by battery life.

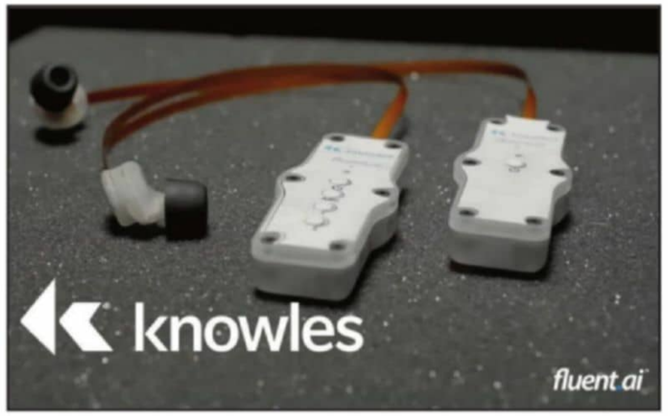
Wireless earbuds on the market today depend on Internet connectivity to sync with devices and respond to user commands. The data processing required from cloud servers and back leads to less battery life, lag in command responses, and limitations regarding when and where the devices can be used.

Hearables currently on the market do not have a truly flexible voice control function. Users require connection to their smartphone voice assistant or a corresponding app for operations or need to tap the devices to carry out commands.

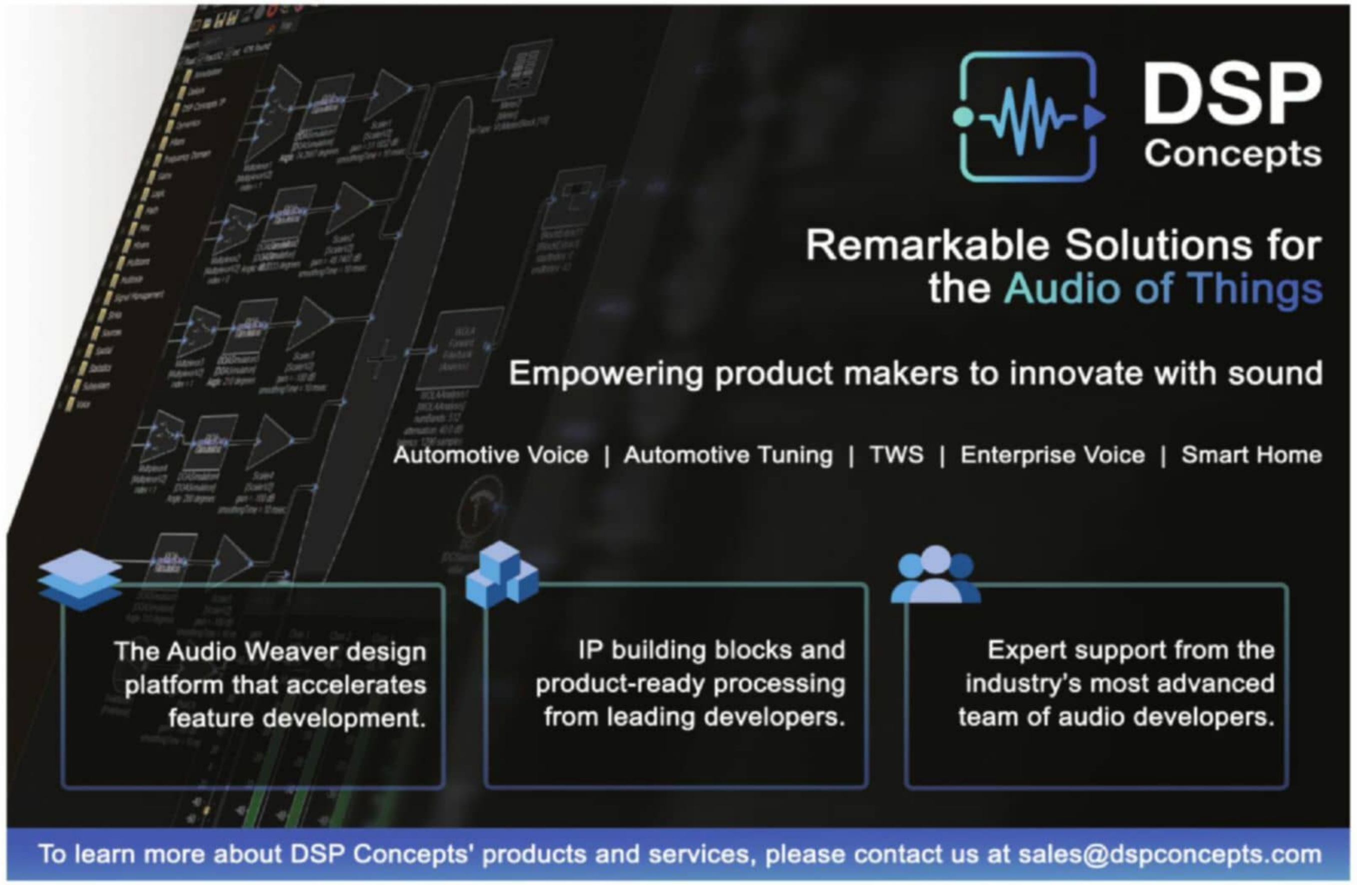
The current technology used in hearables is further restricted by certain languages and has difficulty understanding accents and speech impediments. Usage of the voice command function is also impacted by noisy environments, negatively impacting the user experience.

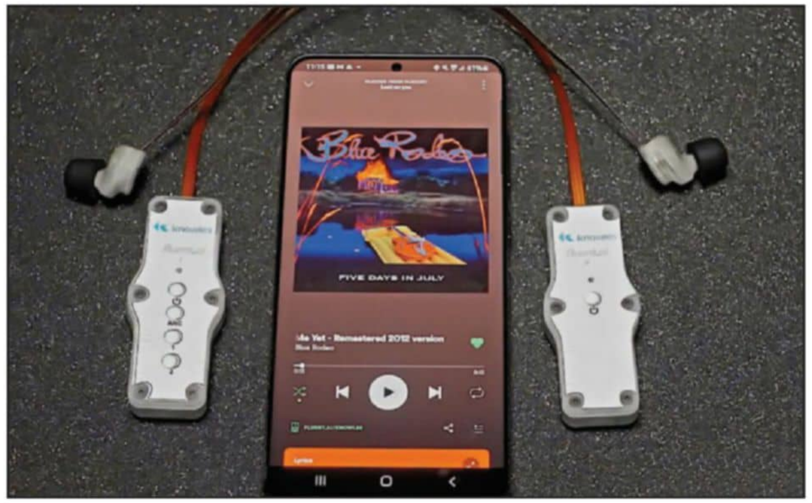
Consumers increasingly expect smart devices to operate with low latency, be energy efficient, and deliver touch-free experiences.

With all of these limitations to overcome, the hearable market has yet to offer the smart technology consumers have come to expect from their devices.



TWS reference design by Fluent.ai and Knowles Corp.





Fluent.ai and Knowles Corp.'s TWS reference design demo

Next Generation of Hearables

With the Internet of Things (IoT), hearables can do more than amplify sound, and thanks to advances in Edge AI and tiny machine learning (TinyML), it's possible to embed the most advanced technology into the smallest of devices (e.g., earbuds).

Knowles Corp., a market leader and global provider of advanced micro-acoustic microphones, speakers, and audio processing, has developed a TWS development kit that enables TWS manufacturers to embed and test advanced audio features. The kit includes the Knowles AISonic Audio Edge Processor IA8201 that is a high-performance, ultra-low-power audio-centric DSP supporting up to four mics and up to 1MB of usable memory, enabling device manufacturers to offer superior algorithms and features. When used in wireless earbuds, the IA8201 processor delivers low-power premium wake-on-voice performance and advanced local commands in addition to a host of front-end audio algorithms for enhanced voice quality (e.g., beam forming, bone conduction sensor fusion, and noise reduction).

Developing the most advanced embedded speech recognition technology available, Fluent.ai is an emerging leader in the voice recognition space. The company's unique speech-to-intent model directly maps a user's speech to their desired action without the need for speech-to-text transcription or a cloud connection. The technology is designed to run fully offline in small footprint, low-power devices and can support any language or accent, providing an intuitive voice user interface for any device or application.

The two companies recently came together to usher in the future of voice-controlled hearables, delivering fully offline and app-free AI-powered voice control for TWS earbuds and other hearable products.

Fluent.ai's patented speech-to-intent technology leverages natural and flexible voice commands to trigger essential TWS actions, including activating noise cancellation, controlling music, accepting or declining calls, and checking battery level, without the need for an accompanying smartphone app, cloud connectivity, or physical tapping of earbuds. The technology resides locally on Knowles AISonic Audio Edge Processor IA8201—which does not use data from the cloud and offers state-of-the art noise cancellation capabilities—and the technology works entirely offline without the need for a Wi-Fi connection.

Moreover, the multilingual and accent agnostic voice recognition technology offers flexible voice commands for a better user experience anytime, anywhere. Fluent.ai's software also offers one of the largest command sets for TWS devices on the market.

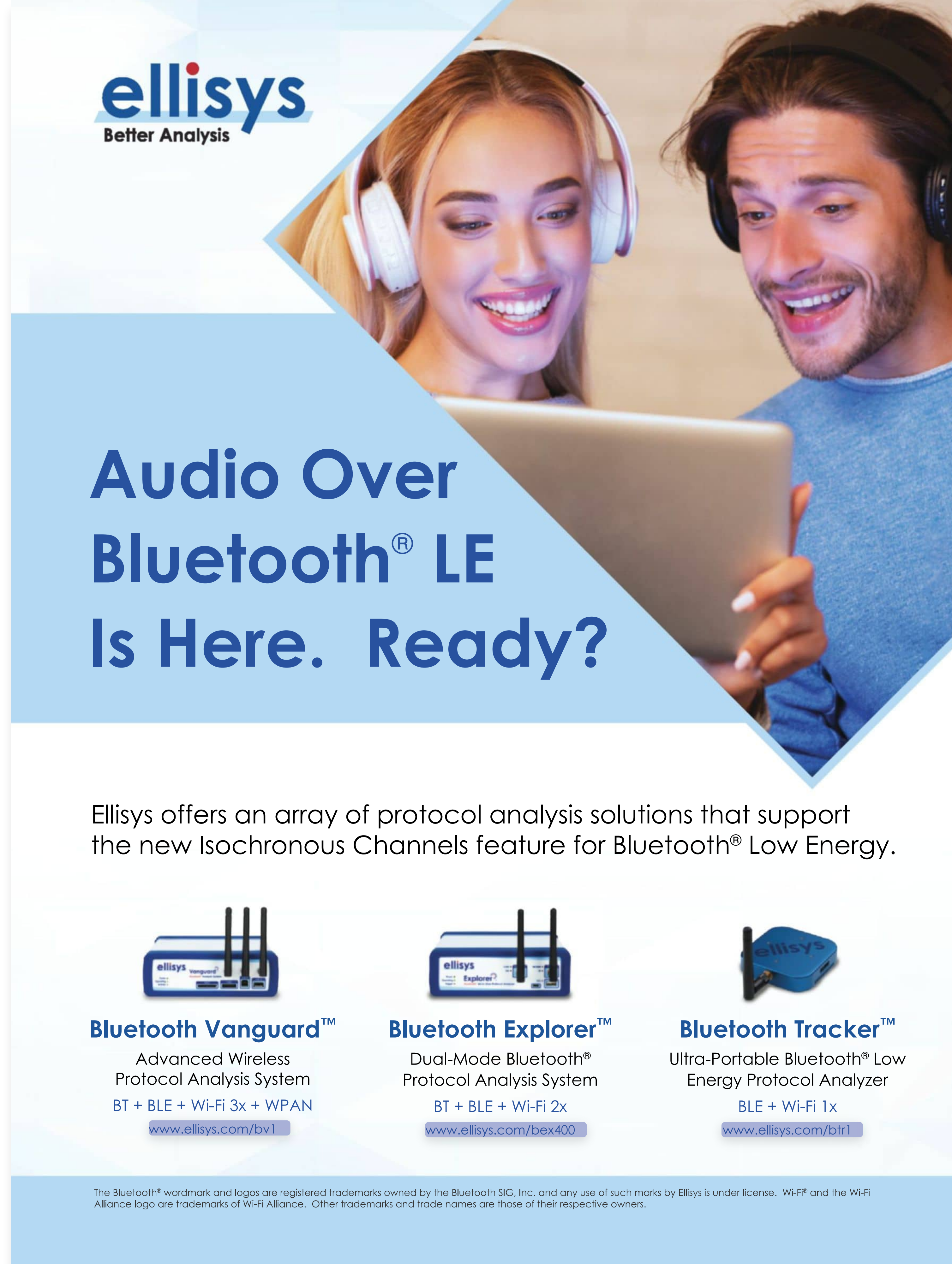
Future of Voice Recognition Technology and Smart Devices

Developers designing voice recognition technology that empowers product makers to deliver noise-robust systems with multilingual support, low latency, and flexible command sets, while reducing costs and shortening the time to market are at the leading edge of future advancements in the hearable, smart device, and voice assistant spaces.

Technologies, such as those mentioned in this article, point toward a future where AI devices will no longer need to depend on an Internet connection, offering consumers the privacy they demand while enhancing the functionality of their smart devices. Aside from the appeal of a more natural interaction with hearables and voice assistants, embedded voice recognition technology demonstrates real-world benefits that incentivize wider adoption. With the processing being performed locally, devices utilizing Edge AI software can be deployed with fewer geographical hurdles as there is no requirement for network infrastructure and language and accent agnostic solutions allow for a truly global adoption.

About the Author

Probal Lala, CEO of Fluent.ai, is a telecom executive turned serial technology entrepreneur and investor with more than 30 years of broad tech leadership experience building high-impact organizations and growing bottom line results. He has been an active Angel Investor for the past 14 years, most recently serving as Chair of the Maple Leaf Angels Capital Corp. Prior to this position, Probal has held a number of executive leadership positions, including being the CEO of About Communications, a national provider of phone and internet services to Canadian businesses; Vice President of Sales at Alcatel-Lucent Canada; General Manager and CTO at Stentor Services, Inc.; and Vice President and General Manager of Internet Operations at Bell Canada.



The Acoustical Challenges of Designing True Wireless Earphones

By
Andrew Bellavia
(Director of Market Development—Performance Audio, Knowles)

Consumer expectations for true wireless earphones are rapidly increasing, raising the bar on new designs. This article from Andrew Bellavia of Knowles Corp. explores the current trends as well as the acoustical challenges and solutions for supporting them.

Several trends in consumer demand, and the technologies enabling them, are converging as the true wireless stereo (TWS) earphone market continues to grow. Use cases are multiplying, with more consumers using wireless earbuds for a diverse and varied set of applications. Music enjoyment is still the number one use case but others growing in popularity include watching videos, listening to podcasts and audiobooks, Internet meetings, conversation enhancement, voice assistants, and gaming. All of these benefit from improved audio quality.

Active Noise Cancellation (ANC) is also becoming more important. In fact it is almost table stakes for any new TWS design [1]. Pre-pandemic, its growth stemmed from people using it for commuting or travel. In the last couple of years, it was also driven by people spending more time working from home and wanting a quiet, focused space. As "normal" life returns and people wear their TWS earphones in varied environments, adaptive or intelligent ANC demand will grow. Adaptive ANC allows appropriate amounts of external sound to be heard according to circumstances, increasing safety and convenience while blocking annoying or even dangerously loud noise.

Amazon Echo (4th Generation) has upward-firing woofer and dual tweeters for rich detailed sound.

Another exciting development is the rise in popularity of hearing personalization (for streaming audio) and enhancement (for speaking with others, especially in noisy environments) [2]. They are the vanguard of technologies that promise to improve the well-being of TWS users through better perceived music quality, reduced listening fatigue during meetings, and less stress in social settings. In combination, these use cases are driving the need for longer battery life and smaller, more comfortable form factors.

Above all these trends, sound quality continues to be at or near the top of the list [3], [4]. Increasingly, improved audio across other consumer products including televisions, automobiles, and smart speakers, such as the Amazon Echo with "upward firing woofer and dual tweeters for rich, detailed sound" [5] is raising expectations for TWS earphones too.

The proliferation of hi-res subscription music services (e.g., Amazon Music HD) further drives this need. The Spotify HiFi pre-launch announcement [6] in February 2021 stated, "High-quality music streaming is consistently one of the most requested new features by our users."

Consumers want hi-res, premium sound through their TWS earphones while supporting the newest use cases, and they want to be able to wear them for long periods of time without worrying about battery life or any compromise to comfort. To achieve this, there are several acoustic design and development challenges. I will discuss these challenges, as well as what OEMs and product designers need to know about creating TWS products that satisfy consumer needs while simplifying design and production.

The Streaming Audio Ecosystem and TWS Earphones

In the early days of TWS earphones, audio quality took a back seat to solving basic challenges. These included improving connectivity and battery life. It wasn't until Apple released the first-generation AirPods in 2016 that the user experience was sufficiently refined to create mass-market appeal. Later came ANC with mainstream brands Sony and Apple introducing products in 2019. At the time, there was little impetus for expending efforts on premium sound because hi-res music streaming wasn't widely available. Some companies, notably Anker/Soundcore with their Liberty 2 Pro model, got ahead of the trend. But by and large, audio quality was not top of mind for brands designing TWS earphones.

All that changed last year with the Spotify Hi-Fi teaser and especially Amazon Music dropping the surcharge for its lossless tier in response to the launch of Apple's equivalent. Now, suddenly, it looked as though TWS earphones were behind the curve in terms of audio quality. While purists at the time opined that it was not yet possible to transmit lossless audio streams via Bluetooth, the truth is that very good codecs exist to take advantage of the improved quality of streaming music. Two of them, LDAC and LHDC, even earned the Japan Audio Society's designation "Hi-Res Audio Wireless."

LDAC is part of the Android Open Source Project, making it widely available on Android phones. Qualcomm, which has the advantage of also supplying chips for many of the world's mobile phones, has also been steadily improving its codecs to deliver hi-res audio through Bluetooth. Most recently, it announced support for CD-quality lossless transmission, ending the last remaining argument about the inherent deficiency of Bluetooth streaming.

Can People Hear the Difference?

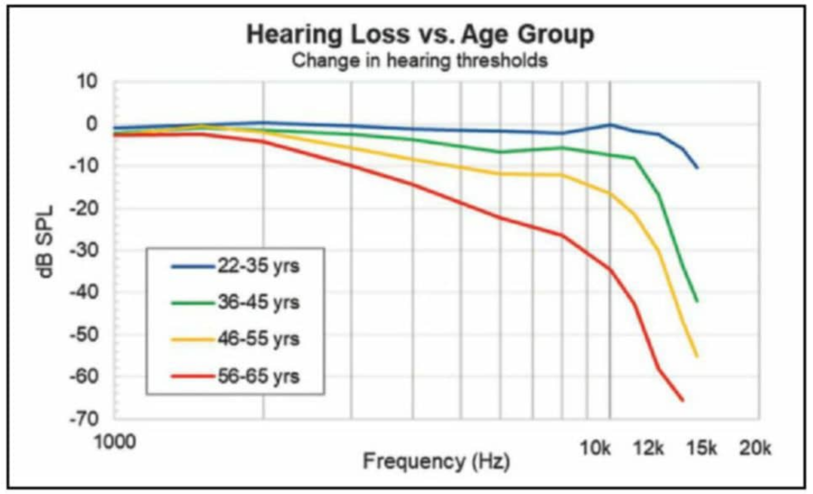
To enjoy hi-res audio, the entire delivery chain has to be optimized for the additional fidelity—including the listener's hearing response. It is a fact of life that most people's treble response rolls off as they age. This is one of the drivers behind the development of over-the-counter hearing aids. However, degradation begins in the treble and therefore affects the perception of music before one notices difficulty in understanding speech.

I once arranged an earphone demo for a hybrid woofer-tweeter speaker system in which the tweeter could be switched on or off. Each person listened and gave feedback. The younger people all loved the two-way driver but the older listeners couldn't hear the difference, likely because of age-related hearing changes.

This potential drag on hi-res audio adoption is countered by the increasing popularity of hearing personalization, not to be confused with hearing enhancement (hearing aid function). Personalization refers to streaming audio, not the microphone response, and is intended to enable more people to get full enjoyment from their music. The increasing demand for premium sound and hi-res audio goes hand-in-hand with hearing personalization. Today, a number of popular brands offer personalization including Skullcandy (with Mimi) [7], Jabra [8], Apple [9], and Anker/Soundcore [10], to name a few.

Format/Codec	Maximum Bit Depth	Sample Rate/ Bandwidth	Maximum Bitrate
Compact Disc	16 bits	44.1kHz/(21)	1411kbps (lossless)
LDAC	32 bits	96kHz/(46)	990kbps
LHDC	24 bits	96kHz/(46)	900kbps
Snapdragon Sound	24 bits	96kHz/(46)	860kbps
aptX Adaptive	24 bits	48kHz/(23)	420kbps
SBC	16 bits	48kHz/(23)	345kbps
AAC	24 bits	44.1kHz/(17)	256kbps (Apple); 320kbps (Android)

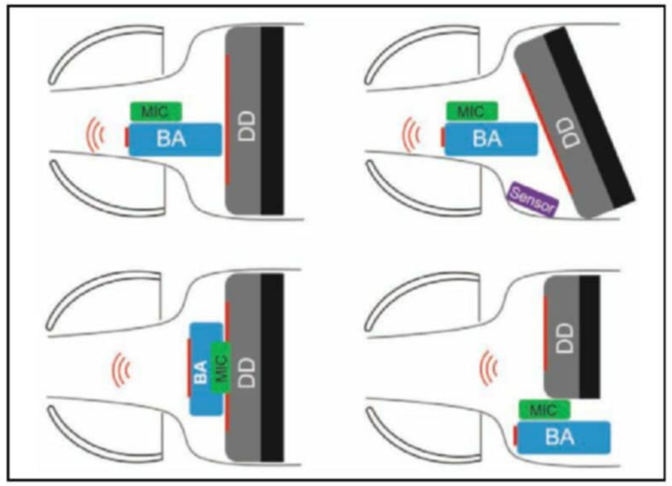
Table 1: Compact Disc, standard, and common Bluetooth codecs are compared.



This graph was derived from "Behavioral Hearing Thresholds Between 0.125 and 20kHz using Depth-Compensated Ear Simulator Calibration," sponsored in part by Hugh Knowles Research Center at Northwestern University.

Meeting the Challenge of Delivering Hi-Res Audio to the Mass-Market

Supporting ANC, hi-res audio, and hearing personalization, while driving longer battery life and improved comfort, increases the challenges for the acoustic designer. ANC requires a speaker that can provide high bass drive levels with low total harmonic distortion (THD). That same speaker must support extended treble for premium music quality, with sufficient overhead to support adding gain for personalization. This is a tall order when simultaneously driving the speaker size down to improve comfort, support room for a bigger battery, additional features, or a smaller form factor.



Here are a few examples showing flexibility of a hybrid driver layout.

Here is the Knowles TWS Development Platform.

A hybrid driver with a dynamic woofer and a balanced armature (BA) tweeter is an ideal solution in which both speakers play to their strengths. Dynamic speakers are excellent for driving the bass necessary for ANC, including a vented opening for occlusion reduction. But designing for optimum bass often comes at the sacrifice of treble performance in the smallest speaker sizes. On the other hand, BAs have always been known for excellent treble performance, high sensitivity and output, and small size. Employing a dedicated BA tweeter allows one to use a true woofer designed for maximum bass in the smallest possible size. The BA tweeter delivers hi-res audio capability with plenty of headroom for hearing personalization.

In addition, separating the woofer and the tweeter creates more flexibility in the layout. A single full-range speaker must be located close to the eartip to maximize treble extension. However, a woofer may be located further back, or mounted off-axis if desired, with the tweeter positioned closer to the ear tip. Since the BA tweeter is very small, this is easy to accomplish even when an ANC microphone and other sensors are also employed.

Avoiding Feedback

Another issue can arise when hearing personalization, especially hearing enhancement, is employed. Boosting the treble in a full-range dynamic speaker can cause feedback through the exterior microphones. To avoid this, the speaker must be acoustically isolated. This can come at a cost of increased size or sub-optimum layout. But a BA tweeter is already a fully enclosed system. Tests have shown that BA tweeters can support 20dB more gain before feedback compared to dynamic speakers.

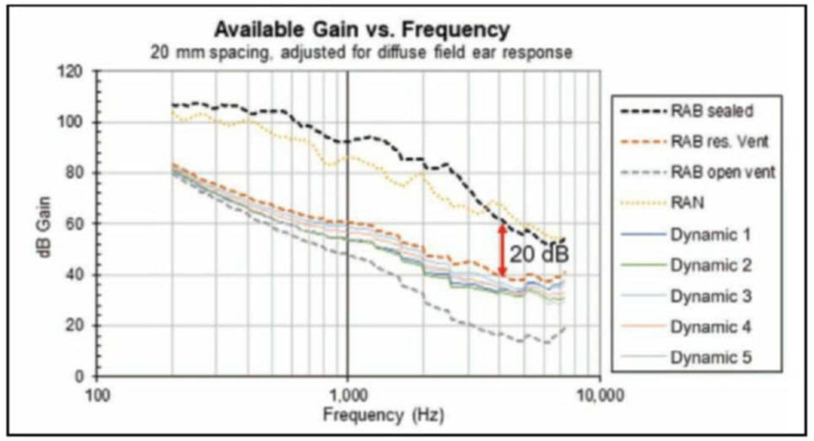
The Knowles TWS Development Platform

Ever at the forefront of audio system innovation, Knowles has created a powerful, configurable platform for developing TWS earphones with premium features [11]. Ultimately, the platform will enable OEMs to decrease time to market by porting pre-tuned hardware and software to a user-defined TWS design.

The kit's pre-configured Knowles earbuds feature proprietary multi-microphone arrays, including a high-performance, low-power Everest digital microphone; voice vibration sensor; and a choice of drivers including a hybrid using a BA tweeter and a small dynamic woofer developed by Ole Wolff in partnership with Knowles [12]. A powerful, flexible IA8201 AISonic audio edge processor enables running advanced voice and AI algorithms form a growing universe of partners in the audio, voice, and hearing spaces.

Conclusion

As the demand for TWS continues to grow, designers face an uphill battle to develop earphones that can check all the boxes of consumer preference: comfort, battery life, and inclusive premium sound. Those that can effectively leverage the latest design techniques, such as integrating balanced armature tweeters in a hybrid configuration, will be able to achieve TWS designs that both delight consumers and reduce manufacturing cost and time to market.



Knowles' measurement of permitted gain before feedback from a microphone placed 20mm behind a speaker mounted in measurement coupler. Sealed BA drivers offer 20dB more gain.

References

[1] Qualcomm, "2021 State of Sound Report," p. 12, www.qualcomm.com/media/documents/files/snapdragon-soundscape-experience-the-state-of-sound-2021-report.pdf

[2] Ibid, p. 18

[3] Ibid, pp. 5, 8

[4] Sonova, "Investor & Analyst Day 2021 Handout," p. 56, September 14, 2021 https://relaunch.sonova.com/cdn/ff/Y30b12YcKlOeU1vNnIz_TMCj5X-6UTv6tDCzKR_fSEg/1631617253/public/2021-09/IR%20Day%202021_Handout_2.pdf

[5] Echo (4th Gen), www.amazon.com

[6] Spotify, "Five Things to Know About Spotify HiFi," February 2021, https://newsroom.spotify.com/2021-02-22/five-things-to-know-about-spotify-hifi

[7] Mimi, "Skullcandy and Mimi Hearing Technologies Partner to Foster Hearing Health," November 2021, www.mimi.io/press/skullcandy-and-mimi-hearing-technologies-partner-to-foster-hearing-health

[8] Jabra, "What is MySound?" www.jabra.com

[9] Apple, "Customize headphone audio levels on your iPhone or iPad," https://support.apple.com/en-us/HT211218

[10] Soundcore, "HearID," https://us.soundcore.com

[11] Knowles, "True Wireless Stereo (TWS) Development Platform," www.knowles.com/tws-development-platform

[12] "Knowles Partners with Ole Wolff to Offer Compact Hybrid Driver for TWS Earphones," *audioXpress*, December 2021, https://audioxpress.com/news/knowles-partners-with-ole-wolff-announce-compact-hybrid-driver-for-tws-earphones

About the Author

Andrew Bellavia is director of market development at Knowles Corp. With a lifelong passion for audio and more than 30 years of experience in the audio field, Andrew is an advocate for innovations in hearing care and sound devices in both his professional and personal endeavors. For nearly 10 years, he has worked to bring Knowles' world-class acoustic solutions and balanced armature technology to market for innovators, leaders, and end users in the hearing health, in-ear monitor, and premium sound markets.

Frequency Response of a Listening Room

This article proves that not only is it possible to measure the frequency response of a listening room, but it is also possible to measure an infinite number of non-identical frequency responses. However, our acoustics expert argues that without standardization of the meaning of the term "frequency response" as applied to rooms, the term is probably not a useful metric from an engineering standpoint.

By
Richard Honeycutt

An audio technician recently asked me, "How do you measure the frequency response of a listening room?" While recognizing that the complex nature of the physics hidden in this seemingly simple question, I nevertheless understood the reasoning behind his asking it. When my primary reading consisted of electronics/audio magazines, I came to the conclusion that knowing the frequency response of any link in the audio chain was a vital part of evaluating its performance. I have since come to realize that the emphasis on frequency response was at least in part a relic of times when that was pretty much the only performance metric that an amateur could measure, and likely the only metric that Joe Public would come close to understanding. "When your only tool is a hammer…"

Consequently, when, as a wannabe acoustician, I performed my first acoustic design of a room—in this case, a small coffee house—I looked far and wide for ways to make sure the frequency response of the room was suitable for live music. However, I searched in vain. By the end of this article, I think you will understand why.

The frequency response of—say—an amplifier or a loudspeaker is usually presented as a graph of the output level plotted against frequency, with the input being a swept sine wave or some other signal containing known levels of all frequencies. Based upon this concept, it would seem that we could apply a known acoustic signal to a room and measure the SPL, thus determining the frequency response of the room. The ointment, however, contains at least five flies:

- Especially in acoustically small rooms, the SPL depends upon the position in the room.
- The nature of the sound source is critical to any quantification of SPL in the sound field: Is it a point source, a line source, a distributed source…?
- The source directivity is also a major factor. Each type of source can have different directivities, depending upon the design and dimensions of the source. A symphony orchestra will not produce the same sound field as a soloist accompanied by piano or guitar; it won't even produce the same sound field as a superb speaker system playing a superb recording of that same orchestra.
- The audience, including the number and average size of the people and how they're dressed and how they're physically arranged, exerts a crucial effect upon the frequency balance heard by an audience member.
- The reverberation signature of the room, including both the sound decay profile and the variation of Reverberation Time (RT) with frequency, is crucial to human perception of the room's frequency balance.

Let's look at each of these factors in turn.

Small Rooms

The Sound Control article published in the April 2019 issue of *audioXpress* discussed the acoustics of small rooms. To summarize the main points of

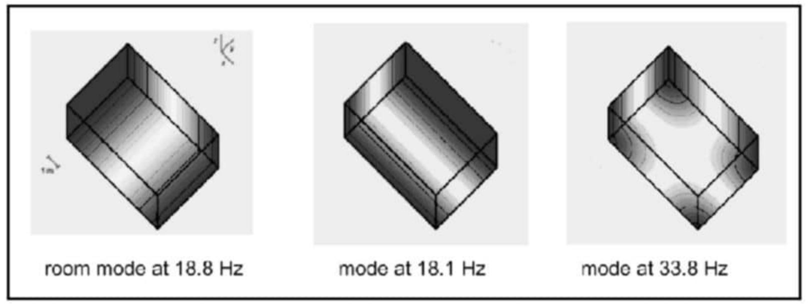


Photo 1: The lowest three modes in this room cause the SPL distribution shown here.

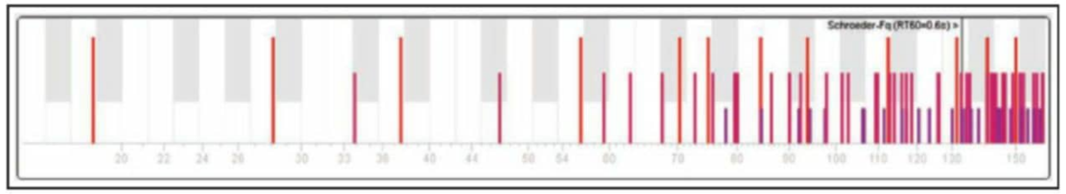


Figure 1: The modes in this room are shown by frequency.

that article, I would mention that in any room, there are sound-reflection paths ("modes") that cause reinforcement of the sound at some frequencies and cancellations at others. Modes can be axial—reflections between parallel surfaces, such as walls or ceiling and floor; tangential—reflections involving all four walls; or oblique—involving both walls and ceiling or floor. The locations in the room where the sound is reinforced or cancelled varies with frequency, as shown in **Photo 1**, for a 20′-wide by 30′-long by 8′-high room. The dark regions correspond to high SPL.

Figure 1 shows the distribution of the modes below 160Hz. Note that as the frequency increases, modes become more closely spaced. The number of modes per one-third octave is critical: according to Oscar Bonello's work [1], the number of modes per one-third octave should increase with frequency. **Figure 2** shows that this room does not meet the Bonello criterion. Decreasing the room's width would bring it into compliance.

Above a frequency called the Schroeder frequency, modes are spaced closely enough that the room response becomes reasonably smooth, and modes are not a problem. The Schroeder frequency is given by:

$$f_s = 2000\sqrt{\frac{RT_{60}}{V}}$$

where RT_{60} = the reverberation time of the room, and V = the volume of the room in cubic meters.

For the 20′ × 30′ × 8′ room with an RT of 0.6 seconds, this frequency is about 133Hz. For a small auditorium measuring 60′ × 80′ × 20′ and having an RT of 1.3 seconds, f_S would be 49 Hz. In larger rooms, f_S is too low for modes to be of concern.

So how can we measure the low-frequency response of a room in order to evaluate the effect of modes? The Sound Control article published in the June 2017 issue of *audioXpress* discusses measurement approaches, with an emphasis on the MATT test devised by Acoustic Sciences, Inc. (www.acousticsciences.com/matt).

Nature of the Sound Source

If the sound source in a listening room is one or more loudspeakers, the directional response of these will affect the frequency response at all points in the room, but in different ways at different listener locations. The frequency response of any real sound source will vary with frequency. In adjusting equalization for a sound system, the technician ideally uses at least three to five microphones located in strategic positions. The outputs of these mics are fed into a multiplexer (see **Photo 2**) that cycles the various mics to feed a Real-Time Analyzer (RTA). In this way, the measured level at each frequency is a time-averaged version of the sounds at each mic position. This procedure could be described as measuring the frequency response of the room, but its results do not correspond exactly to what any individual listener hears.

A source such as a symphony orchestra or a choir really comprises a large group of point sources with differing spectral contents and directivities. The same can be said for a pipe organ. Much study has been devoted to finding ways of simulating the sound field of such sources, beginning with the Bell Labs demonstration (described in the "Symposium

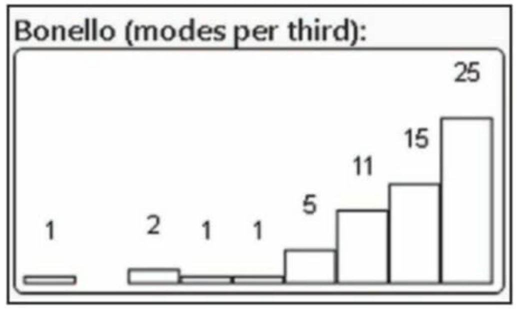


Figure 2: Modes per one-third octave should increase with frequency.

Photo 2: The Mirideo MUX-PRO multiplexer time-averages signals from up to four mics.

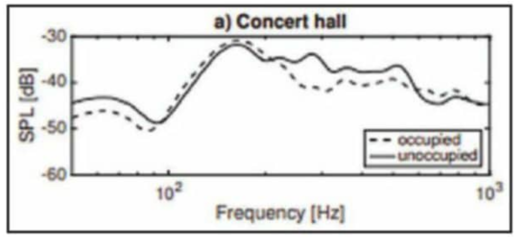


Figure 3: Diffraction above the audience affects the frequency response at a listener seat.

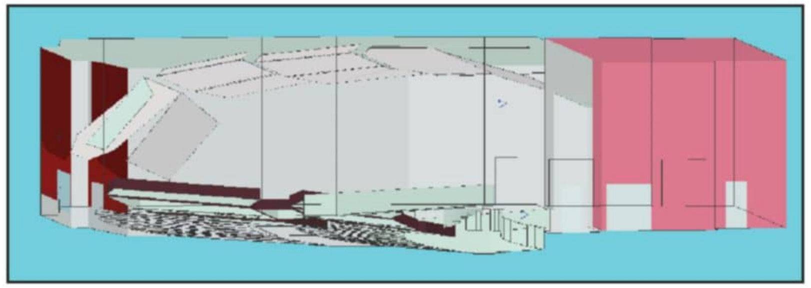


Photo 3: This is an EASE computer model of a 171,000-cubic-foot high school auditorium.

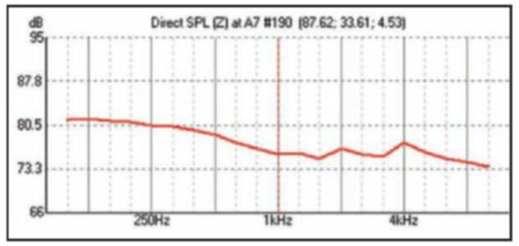


Figure 4: The direct SPL at location X=87.62', Y=33.61', Z=4.53' is higher at low frequencies and irregular at high frequencies.

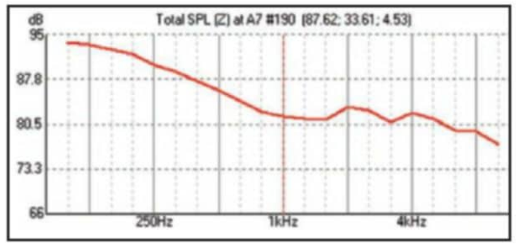


Figure 5: The total SPL (including reverberation) is about 13dB higher than the direct SPL, and has greater variation with frequency.

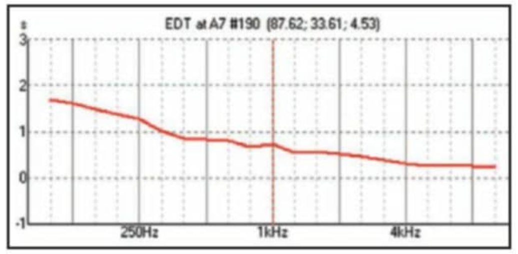


Figure 6: The EDT decreases from 1.75s at 100Hz to 0.25s at 8kHz.

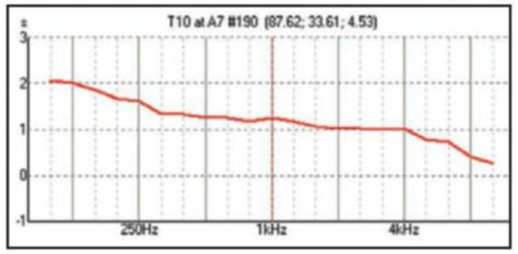


Figure 7: The RT based upon the first 10ms of decay also decreases as frequency increases, but is somewhat higher than the EDT, especially at low frequencies.

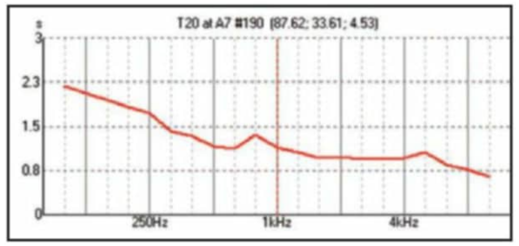


Figure 8: The RT based upon the first 20ms of decay also decreases as frequency increases, but is somewhat higher than T10.

on Audio Perspective") in which three microphones picked up the sound of the Philadelphia Symphony Orchestra playing in the Academy of Science Hall in Philadelphia and reproduced through a three-channel sound system incorporating three Shearer horns in Constitution Hall in Washington, D. C., "with a fidelity, depth, and spatial effect that effectively created the illusion of the orchestra's presence behind the stage curtain" [2].

While this demonstration was phenomenal for its time, I think it likely that an audience of 21st-Century sound engineers would find a modern surround-sound system's performance more realistic. And even the best surround-sound system cannot truly recreate the exact sound field of a performance held in another room.

The Effect of the Audience

At a given listener seat, the effect of the audience upon the perceived frequency spectrum manifests in four ways:

- A listener seated behind a taller person may have some high frequencies blocked by the person in front.
- Rows of people create a sort of diffraction grating, which creates the "seat-dip effect," a dip in frequency response typically in the 200Hz to 500Hz range accompanied by a small boost between 600Hz and 800Hz (**Figure 3**). This effect varies depending upon seat spacing and height.
- The occupancy (number of people) in any room affects the RT, since each adult adds about 3 sabins to 4 sabins, particularly in the mid-frequency range. The effect of RT upon perceived frequency balance is discussed below.
- In winter, audience members may add more acoustical absorption due to heavy clothing.

The Effect of RT

The sound perceived by a listener depends greatly upon the frequency balance of the RT of the room. **Photo 3** shows a computer model of a high-school auditorium. The EASE modeling software offers "frequency responses" of the direct sound and the total sound (**Figure 4**), including reverberation (**Figure 5**). The frequency balance of the RT can be calculated based upon the Early Decay Time (EDT–six times the time required for the first 10dB of decay), T10, the first 10ms of decay; T20, the first 20ms of decay; or T30, the first 30ms of decay. **Figures 6-9** show the frequency dependence of RT upon frequency.

Rooms having a 125Hz RT that is more than 140% of the midband (500Hz) value will sound boomy; if the 125Hz RT is less than the midband RT, the room will sound thin. Likewise, the human ear/brain expects RT to decrease with frequency above about 1kHz, and if it does not, the room will sound harsh. But if RT decreases too rapidly as frequency increases, the room will sound dull. These effects are difficult to quantify in any form resembling an amplifier's frequency response curve.

Given that the different RT measurements (EDT, T10, T20, and T30) vary somewhat from each other, the room's "frequency response" will also vary with time, so which value of RT is most indicative of a listener's perception? People's judgment of the reverberance of a room largely depends upon EDT. Although there may not be any studies to document it, one would logically assume that people's perception of the frequency balance of a room is also closely related to the frequency balance of the EDT.

Another acoustical parameter related to both RT and thus to perceived frequency response is the sound strength, G. This is a decibel ratio of the SPL at a certain distance from the source in the room, to the SPL the source would produce at the same distance in an anechoic environment such as out-of-doors. An argument can be made that G corresponds to the frequency response of a room; however, the calculation of G does not include factors such as the seat-dip effect, and the time dependence of RT perception (**Figure 10**).

Summary

The concept of frequency response seems simple, and for amplifiers, it is. It is more complex when applied to loudspeakers because of the variation of directivity with frequency (not to mention motional nonlinearity. A physicist might point out that since a speaker has more degrees of freedom than an amplifier, determining frequency response becomes more complex. Likewise, a listening room has many more degrees of freedom than even a loudspeaker, so the concept of frequency response is fraught with even more complexities.

We might summarize the situation by saying that not only is it possible to measure the frequency response of a listening room, but it is possible to measure an infinite number of non-identical frequency responses. So without standardization of the meaning of the term "frequency response" as applied to rooms, that term is probably not a useful metric from an engineering standpoint.

And the above is true even without our mentioning that a given frequency response can correspond to an infinite number of phase responses in a non-minimum-phase system, such as an auditorium [3]!

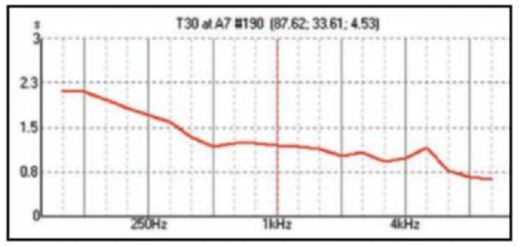


Figure 9: The RT based upon the first 30ms of decay is similar to T20.

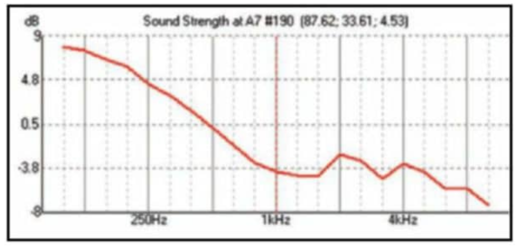


Figure 10: The sound strength G is affected by RT, and thus is also related to frequency response.

About the Author

Dr. Richard Honeycutt fell in love with acoustics after his father brought home a copy of Leo Beranek's landmark text on the subject when Richard was in the ninth grade. Richard is a member of the North Carolina chapter of the Acoustical Society of America. Richard has his own business involving musical instruments and sound systems. He has been an active acoustics consultant since he received his PhD in electroacoustics from the Union Institute in 2004. Richard's work includes architectural acoustics, sound system design, and community noise analysis.

References

[1] O. J. Bonello, "A New Criterion for the Distribution of Normal Room Modes," *Journal of the Audio Engineering Society*, Volume 29, Section 9, pp. 597-606, September 1981.

[2] H. Fletcher, Symposium on Auditory Perspective," originally printed in 1934, and reprinted by the Audio Engineering Society, www.aes.org/aeshc/docs/bell.labs/auditoryperspective.pdf

[3] Thanks to Pat Brown of SynAudCon for pointing this out!

Resources

R. Honeycutt, "Wave Behavior of Sound in Rooms," *audioXpress*, April 2019.

R. Honeycutt, "Assessing Audio Quality Using the MATT Test, *audioXpress*, June 2017.

H. Tahvanainen and T. Lokki, "About the Seat-Dip Effect in the Presence of Audience," Research Gate, July 2018, www.researchgate.net/publication/326461386_ABOUT_THE_SEAT-DIP_EFFECT)

GaN Technology in Audio Power Amplification

By
Stuart Yaniger
Photos by
Cynthia Wenslow

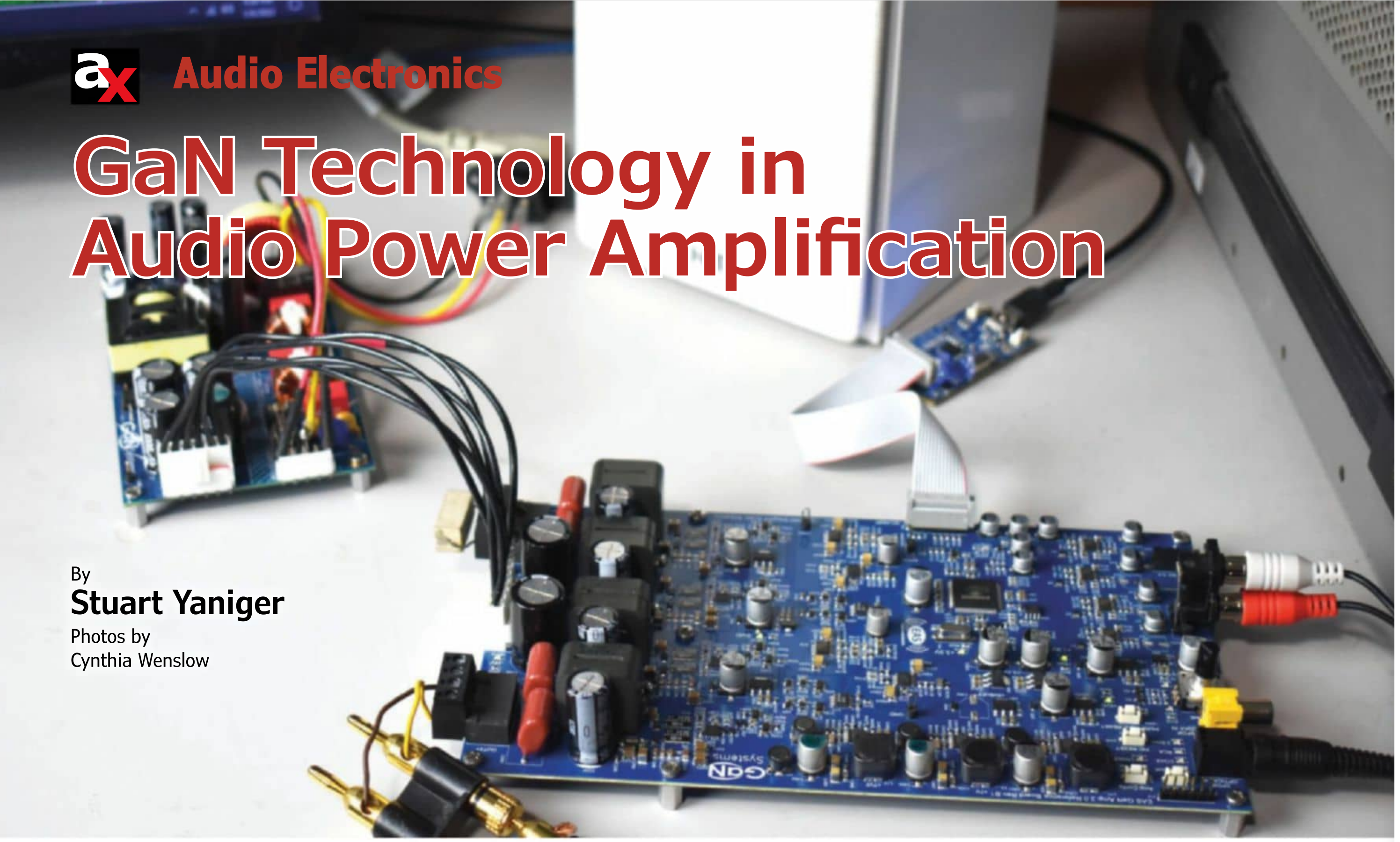


Photo 1: The amplifier EVAL Board from GaN Systems is an efficient two-channel, 200W per channel (8Ω) Class-D audio amplifier and companion 400W continuous power audio-grade switched-mode power supply.

As he reviews a Class-D amplifier and switched-mode power supply (SMPS) evaluation kit from GaN Systems (**Photo 1**), Stuart Yaniger details the benefits for audio applications and shares his experiences with gallium nitride (GaN) power semiconductors. His set of measurements of the evaluation board enables us to better understand the performance potential with and without feedback loops from the output stage to the controller.

In a philosophical sense, there are different levels of technology—most commonly, we think of fundamental and applied technologies. An example of the former would be the invention of transistors at Bell Labs in the 1940s. The driving force behind the fundamental advance was telecommunications, a much wider market than audio, but the availability of devices using the new technology transformed the consumer electronics markets. Once the concept of controlling electron flow in a semiconductor was established, it was a matter of application to use this technology in places such as audio, and over the years, audio-specific devices were created based on the fundamental telecom technologies. Audio was not the driver of technological change, but was the lucky recipient of its benefits.

Analogously, integrated circuits (ICs), computer networking, optoelectronics, lasers, micromechanical devices, and nanotechnology were not created for audio, but had a rapid uptake once they were commercially available because of the huge advantages in performance, efficiency, and reliability. Of course these days, there's no shortage of audio use of exotic materials, such as graphene, carbon nanotubes, and the like.

While some technological adoption has been more for buzzword and gimmick value, gallium nitride (GaN) semiconductor devices are the real deal and have great promise for real performance gains in consumer and professional audio. In this article, I'll briefly review what GaN is all about, why it's important in general and to audio specifically, and discuss a couple of real-world examples.

GaN In a Nutshell

GaN is a wide bandgap semiconductor (3.4eV compared to 1.1eV for silicon and 0.67eV for germanium) originally used in optoelectronics—its wide bandgap means that the photons emitted are high energy, and thus short wavelength. GaN's adoption (as with the similar gallium arsenide) was key to achieving efficient blue LEDs. A major difference from silicon is that the gallium-nitrogen bond is polar, and structures can be formed with one dimension (parallel to the Ga-N bond) being very different than the other two. The wider bandgap means that GaN devices can have higher breakdown voltages and withstand higher operating temperatures as well. Additionally, GaN is a direct bandgap semiconductor, that is, one where the top of the valence band and the bottom of the conduction band have the same momentum. This is key to its utility in optoelectronics, because promotion of electrons from valence to conduction bands only requires absorption of a photon of energy equal to or greater than the bandgap instead of needing the photons to interact with lattice phonons, as is the case for indirect bandgap semiconductors.

The unique structure of GaN crystals allows the creation of thin layers having exceptionally high electron mobility, defined as the ratio of electron velocity to applied field. So high-mobility electrons can be accelerated to high velocity, which implies high attainable device speed and efficiency. Because the electrons are constrained to a very thin layer in the vertical direction but have high mobility in the two horizontal directions, the electrons act as though they were particles in a gas (i.e., almost entirely free to move with low scattering). Thus, this layer is called a "two-dimensional electron gas," or 2DEG. This 2DEG is key to the efficiency and speed performance of GaN devices.

So we effectively have a sheet of highly mobile electrons. The conductivity needs to be switchable, in the manner of a gate in a conventional MOSFET, so a depletion region is created. A gate electrode is then used to increase electron mobility and density in that region for device turn-on. Compared to silicon, GaN is a more expensive material, and requires different (read: high investment) capital equipment for device fabrication. To overcome this issue, manufacturers developed processes to allow GaN fabrication on a silicon substrate. A typical construction for such a GaN High-Electron-Mobility Transistor (HEMT) is shown in **Figure 1**. The gate can either be isolated, with field used to turn on the conductivity, or fabricated as a heterojunction, with differing band-gaps used to create depletion under zero gate voltage conditions.

Besides increased electron mobility, there are several other advantages for GaN HEMTs' speed. First, the gate resistance itself is quite low since the electrodes can be fabricated from metal rather than, for example, silicon nitride. Second, the gate threshold voltage (V_{th}) is significantly lower (roughly half or less) than that of silicon-based MOSFETs. For GaN Systems' devices, the V_{th} is about 1.5V, with full turn-on by 5V V_{gs}. Third, unlike conventional MOSFETs, there is no intrinsic body diode. This means that the reverse recovery charge is vastly reduced to near zero. Putting numbers on that, Q_{RR} for a silicon MOSFET will be somewhere north of 1000nC and vary with charge/discharge rate,

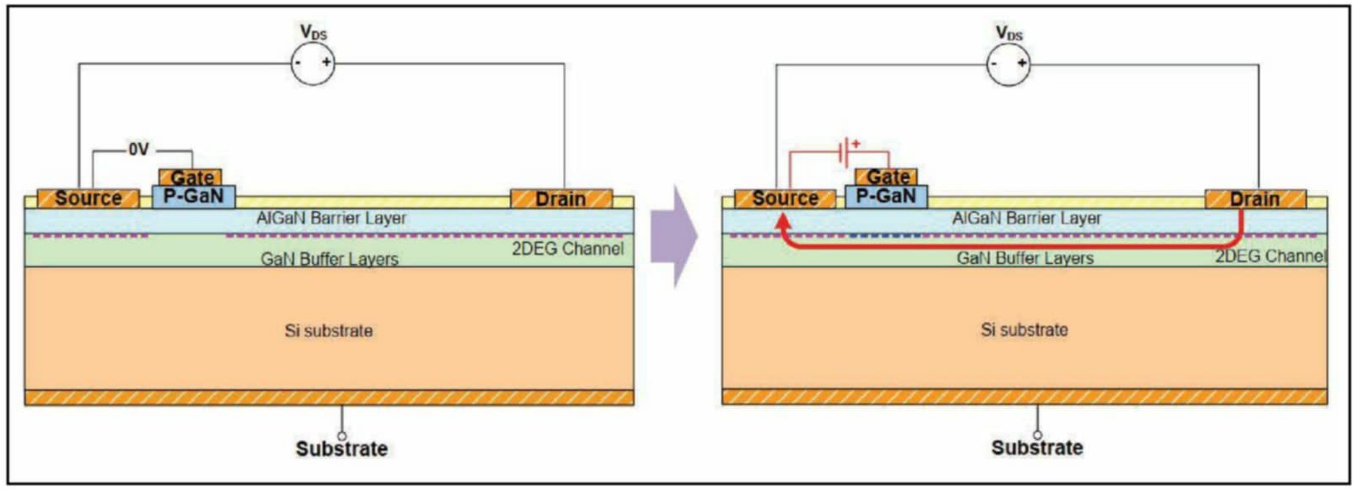


Figure 1: This diagram shows the structure of a typical GaN power HEMT with and without turn-on bias. (Courtesy of GaN Systems)

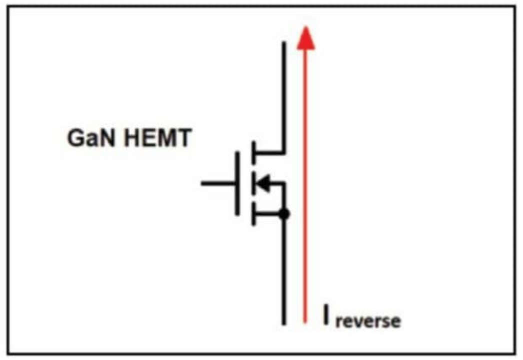


Figure 2: Reverse current flows through the GaN HEMT's conduction channel without a diode.

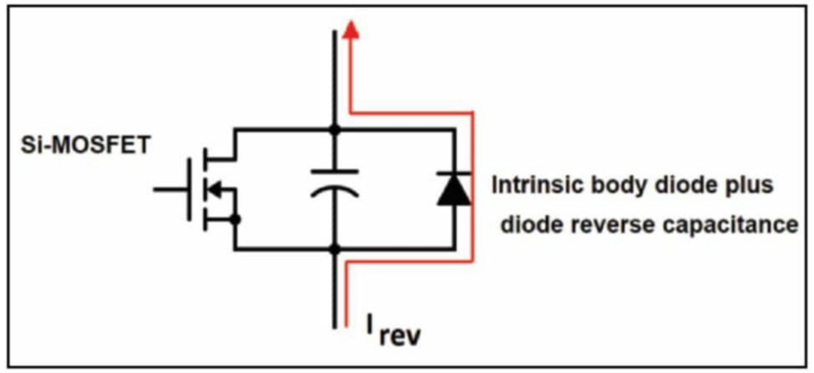
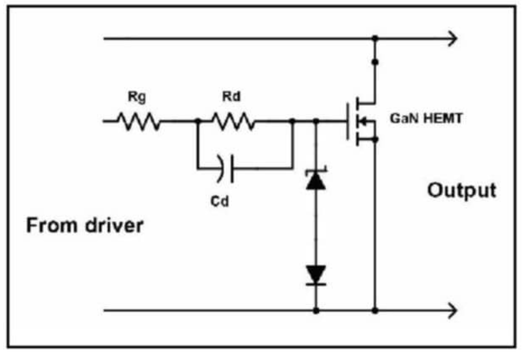


Figure 3: Reverse current flows through the intrinsic body diode of a conventional MOSFET.

Figure 4: Conventional MOSFET drivers can be used with GaN Systems' HEMTs by using the EZ Drive circuit.



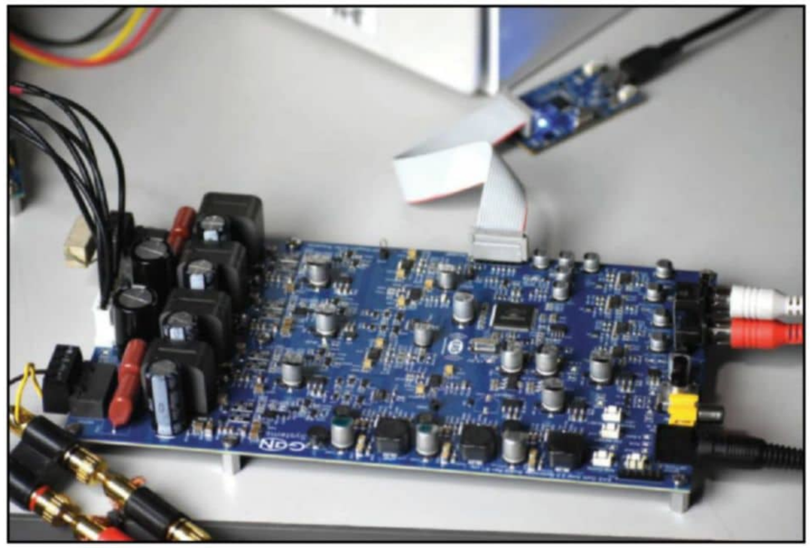
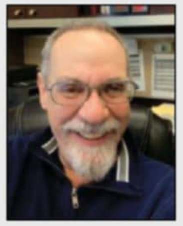


Photo 2: Before starting to measure the board's capabilities, I hooked it up in my lab system for an afternoon, just to establish a baseline.

About the Author



Stuart Yaniger has been designing and building audio equipment for nearly half a century, and currently runs a technology consulting agency in western New York. His professional research interests have spanned theoretical physics, electronics, chemistry, spectroscopy, aerospace, biology, and sensory science. One day he will figure out what he would like to be when he grows up.

whereas a GaN Systems HEMT will have a Q_{OSS} of typically 60nC. These translate to faster and cleaner switching, which allows the reduction in deadtime and consequent reduction in distortion.

The fundamental operation and structure of the GaN HEMT allows reverse conduction through the body of the structure (**Figure 2**), in contrast to MOSFETs (**Figure 3**). This means an external diode path is also not needed, simplifying design. Because of the lower parasitic charges, the switching currents are also much cleaner, with far less ringing. This means that electromagnetic interference (EMI) suppression is much simpler.

The GaN Systems EVAL Boards

GaN Systems, a Canadian-based fabless semiconductor manufacturer, is one of the top suppliers of GaN power devices for automotive, power, and consumer electronics. To assist designers and OEMs in creating audio amplifiers using GaN technology, GaN Systems sells evaluation boards with a reference amplifier design and an accompanying power supply (Photo 1). These units were designed for GaN Systems by Skip Taylor at Elegant Audio Solutions (EAS), an Austin, TX-based design and engineering firm. Both the power supply and audio power amp circuits can be easily adapted for specific applications but are meant to highlight the preferred design practices to achieve a good trade-off between cost and performance, while optimizing the efficiency and cool-running advantages of their GaN devices.

The GaN Systems reference design power supply is available with either a 12V DC input for designs targeted to automotive or 85V AC to 264V AC input for home audio. The supply outputs ±32V DC at up to 400W at efficiencies over 90% at full power out. It incorporates power factor control functionality as well, based on the NCP1654 PFC controller. The power circuit is configured as a half-bridge LLC resonant circuit, and both the PFC and power circuit use GaN System GS66506T 650V-rated HEMTs as power devices.

One interesting aspect of this supply is the use of a simple modification to the circuit driving the HEMTs. Since GaN HEMTs have different drive requirements (lower threshold voltage) than conventional MOSFETs, the designer has a choice between using controllers/drivers made for MOSFET circuits or controllers/drivers specifically made for GaN devices. The former choice is often made because of design features incorporated into the "older" technology and lower costs because of the scale. In order to compatibilize a MOSFET driver with a GaN device, some GaN devices have drivers built in, which are in turn driven by the external MOSFET controller/driver. This of course adds cost and redundancy.

Instead, the GaN Systems power supply uses what they call an "EZ Drive," which allows the use of a MOSFET controller/driver (e.g., an IRS2795), without having a second driver built into the GaN device. The EZ Drive circuit is remarkably simple and elegant (**Figure 4**), and consists of two resistors, a capacitor, and a Zener clamp.

The amplifier board itself is amazingly flexible. It has both analog (RCA and 3.5mm phone) and digital (SPDIF optical and coaxial) inputs, selectable by a push-button switch. There's also a rotary optical encoder for volume control. The power stages use a full ground-referenced bridged output circuit driven by a pair of half-bridge controllers (TI LM5113) designed specifically for GaN circuits.

The output devices are GS61008P, which are rated at 100V and 90A. GaN Systems chose those devices instead of the 650V-rated HEMTs used in the power supply because of slightly easier drive and lower conduction losses. Output terminals are simple screw connectors. One interesting detail is the use of a Kelvin connection from the driver to the sources of the HEMTs, rather than connecting them all at the same point. This can result in a few decibels lower distortion, which is electrically significant when we get down to ultralow distortion levels. It's a nice bit of attention to detail.

A very interesting feature of the GaN Systems EVAL board is the "open loop-closed loop" option, selectable by a push-button switch on the board. In the closed-loop mode, the audio signals from each side of the bridged outputs are fed into a differential amplifier, whose output is then used as

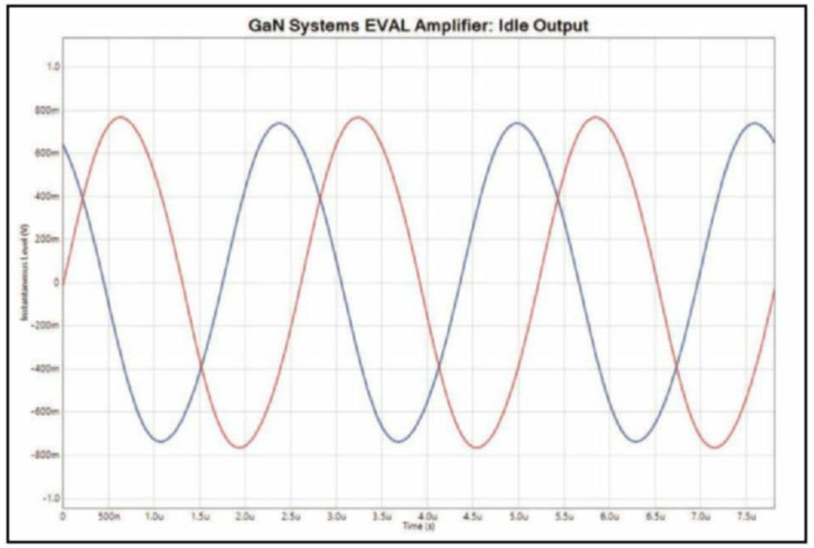


Figure 5: The output of the EVAL board shows the switching frequency of 384kHz with left and right channels in quadrature.

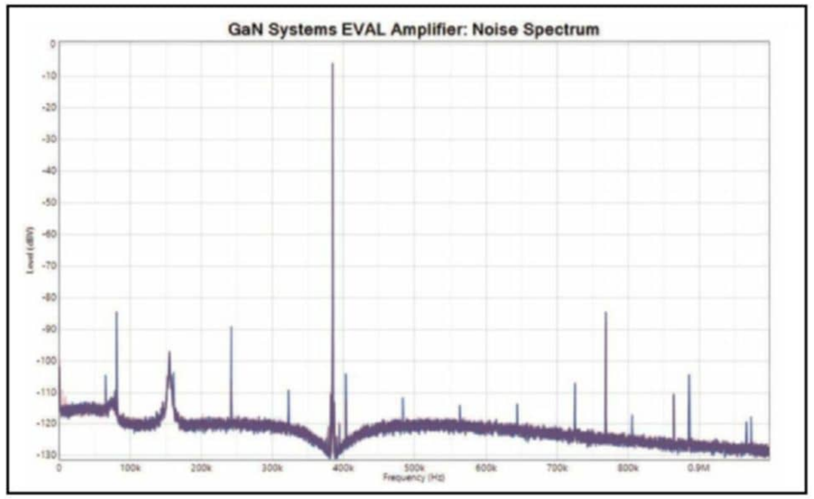


Figure 6: The noise spectrum of the EVAL amplifier is dominated by the switching frequency. All other ultrasonic spurs are below -85dBV.

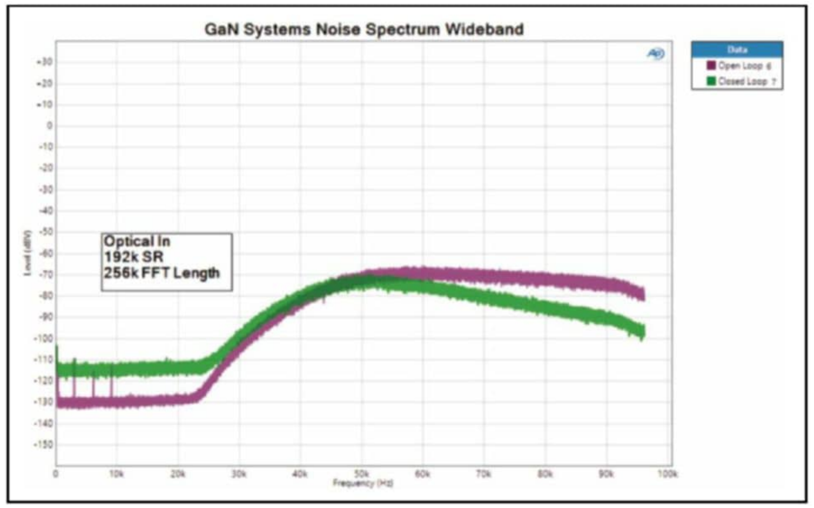


Figure 7: The noise floor in open-loop mode is 15dB quieter than closed loop over the audible range.

Resources

GaN Systems EVAL Board Manual, https://gansystems.com/wp-content/uploads/2020/05/GS-EVB-AUD-xxx1-GS_Technical-Manual_Rev-200526.pdf

GaN in Class D Audio, https://gansystems.com/class-d-audio-transistor-comparison

"An Introduction to GaN Enhancement-Mode HEMTs," GaN Systems, July 2021, https://gansystems.com/wp-content/uploads/2021/07/GN001_An-Introduction-to-GaN-E-HEMTs-210720.pdf

B. Putzeys, "This Thing We Have About Hysteresis Distortion," Purifi Audio, https://purifi-audio.com/2020/04/28/dist

"D2-3(S) Audio SOC Intelligent Digital Amplifier and Sound Processor," Renesas Electronics D2Audio Datasheet, Renesas, April 2019, www.renesas.com/us/en/document/dst/d2-926xx-d2-3s-audio-soc-datasheet?r=502821

a single-ended feedback signal to an earlier stage. In open-loop mode, that overall feedback loop is removed and gain re-adjusted in the signal flow. This provides a nice demonstration of the performance of the GaN output stage on its own.

Note that the open-loop mode only removes the overall feedback loop; local feedback around the op-amps used in the signal flow remains in place. I should also note that feedback is taken from the output HEMTs directly, before the output filter inductor, so the output inductor always stays outside the feedback loop. It has been claimed (see Resources) that this may have audible consequences, but to date, I am unaware of any rigorous double-blind comparison to support this.

Before starting to fool around with the board's capabilities, I hooked it up in my lab system for an afternoon, just to establish a baseline (**Photo 2**). I drove the EVAL board from the optical SPDIF input. In closed-loop mode, it sounded the same as the amp it replaced (which is good!). Switching to open loop, the gain actually dropped, which suggests something was going on in the audio processing beyond a simple loop-breaking. And once levels were matched, it did sound slightly different than the closed-loop mode on my reference amps. And different in a way that was familiar—one of my fun activities is designing and building tube amps, and darned if it didn't sound just like one of them!

This got me looking at some basic performance attributes. For an uninteresting set of reasons, I was unable to use the Audio Precision source impedance utility, so I measured it manually by changing the load resistance and measuring the change in output voltage. From this, I could calculate the source impedance, which for the closed-loop mode was about 0.15Ω at 1kHz. In open-loop mode, the source impedance rose to nearly 1Ω. At 20kHz, the differences were even more significant, with closed loop measuring about 1Ω and open loop nearly 2.5Ω. These impedances will interact with a loudspeaker load (which varies greatly with frequency for most speakers) causing a frequency response difference. And that explains my subjective observation about the similarity to my tube amp. So, the open-loop mode can be an excellent way of dialing in the so-called "tube sound!" For the rest of my evaluation, I kept the amplifier in closed-loop mode.

Putting the amp on my lab bench, I first looked at the output with no signal applied. The oscilloscope trace (**Figure 5**) shows that the oscillation frequency is on-spec at 384kHz. The level at the output terminals is about 500mV, which is par for amplifiers in this class, and absolutely insignificant

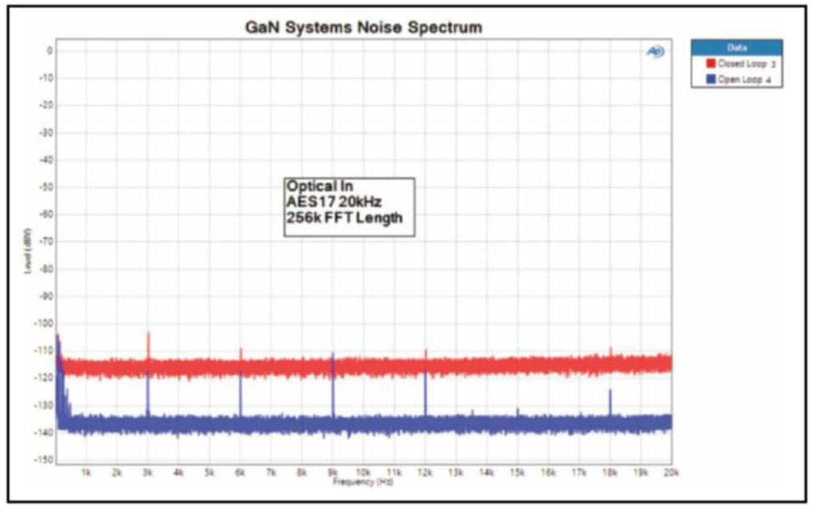


Figure 8: The noise floors in the audible range of open- and closed-loop mode are flat and low, with an advantage to open loop.

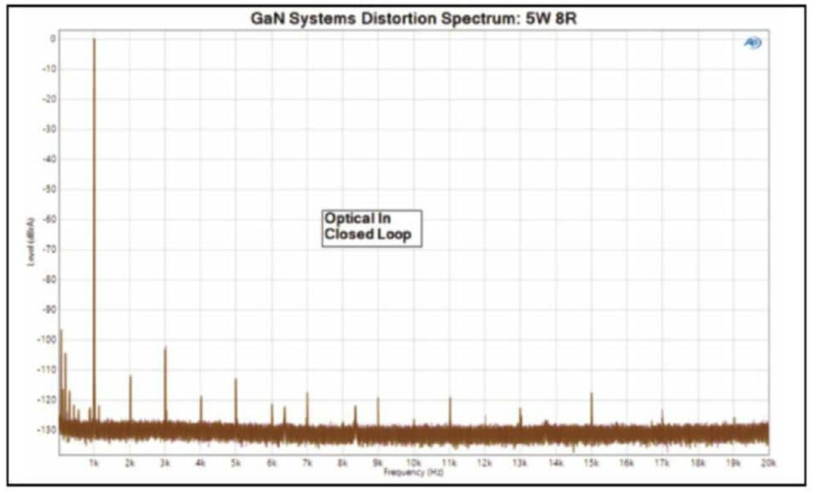


Figure 9: The distortion spectrum at 5W into an 8R load is extremely clean, with the worst harmonic (third) being over -105dB, down from the fundamental.

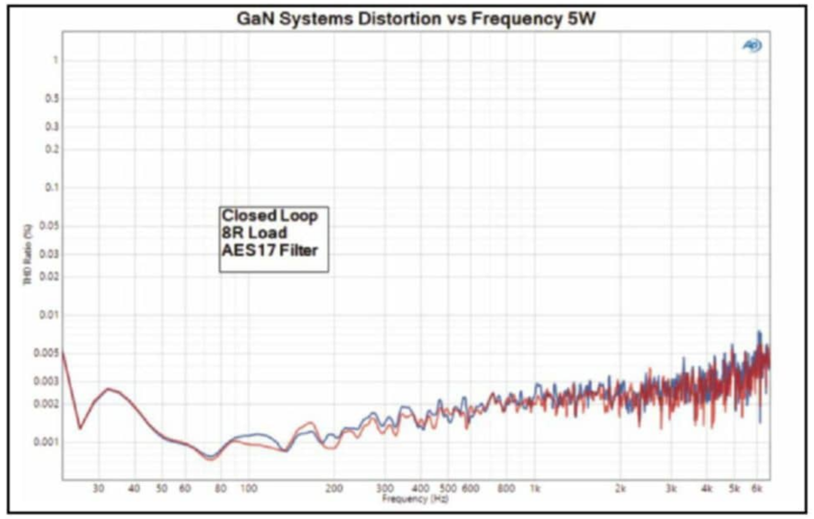


Figure 10: The distortion versus frequency is low and relatively flat.

with a speaker load. Interestingly, the signals for the two channels are roughly in quadrature rather than phase-locked. When asked about this, Skip Taylor of EAS responded that this is intentional and can help with reduction of EMI. I do not have the means to test this, but it seems plausible. The wide-band noise spectrum (**Figure 6**) confirms that the switching frequency dominates, with all other ultrasonic spikes below -85dBV. I asked about this choice of frequency, since GaN HEMTs are capable of switching at much higher frequencies than this. EAS responded that, in their experience, optimal performance is achieved at 400kHz or so, and that going faster brings no real-world benefit.

The feedback topology mentioned earlier has some interesting consequences regarding noise. **Figure 7** shows the noise spectrum with no input in open- and closed-loop modes. In both cases, noise shaping pushes it out of the audible range and into the inaudible ultrasonic. But note that the open loop shows about a 15dB lower noise floor in the audible range (shown in more detail in **Figure 8**). This is likely due to the output noise of the differential amplifier in the feedback loop being amplified through the gain blocks preceding the driver chip. The closed-loop setting is by no means noisy, we're talking about the difference between inaudible and really, really inaudible. Nonetheless, this is at least a theoretical and measurable advantage for open-loop operation under this particular feedback topology.

A quick check of power output showed that, with the power supply accompanying the EVAL board, clipping began at 150W into an 8R dummy load, both channels driven. Reducing the load to 4R, the amp's power actually declined to 120W, suggesting that the limitation here is the power supply.

Figure 9 shows the distortion spectrum at 5W into an 8R dummy load. The dominant component is the third harmonic at less than -105dB down. This is excellent performance. Equally excellent is the relatively low frequency dependence of the distortion (**Figure 10**). Because of the AES17 filter, the measurements are only valid to slightly under 7kHz (the dominant third harmonic will exceed the measurement bandwidth), but the very slow rise with frequency is evident. It's possible that even here, the dominant source could be noise, given the very low levels involved.

One question that arises again and again in discussions of amplification, and especially the newer Class-D technologies, is the effect of real-world speaker loads on performance. One of the speakers I regularly use is an old NHT two-way mini monitor. Its measured impedance is shown in **Figure 11**. I have used this impedance to create a reactive dummy load, but in my discussions with EAS, they felt that an actual speaker acted differently than a pile of coils, capacitors, and resistors. So I bit the bullet, plugged my ears, and ran some sweep measurements at 6.3V (corresponding to 5W into 8R). **Figure 12** shows the frequency response of the EVAL into the actual physical speaker. The interaction between the output impedance and the reactive speaker load causes about a 0.25dB lift in the top two octaves. This might possibly be audible with direct rapid-switch ABX-style comparison and a pair of really sensitive ears, but is skirting human performance thresholds, so is unlikely to be a factor in real-world use.

The distortion versus frequency sweep with the speaker load is shown in **Figure 13**. There are some peaks and dips correlating with the impedance changes in the speaker, but the distortion is still very low and generally flat. This is excellent performance for the intended application: driving loudspeakers.

With the basic performance verified, it was time to move on to some of the DSP capabilities of this EVAL board. The audio processing is handled by a Renesas D2Audio 92633 Digital Control Processor (see Resources). The processor is programmable via a SCAMP7 dongle (available separately from the EVAL board), which attaches to a port on the board and connects to a computer via USB. Audio Canvas III software from Intersil (not to be confused with the music production software of the same name) allows the designer to program the board's firmware. Its capabilities include hardware adjustment, signal flow, signal conditioning, and DSP. This software is about 10 years old now, so the user interface feels a bit clunky; nonetheless, it's pretty easy to use and reasonably intuitive.

With the dongle attached and the drivers installed, you can choose the EVAL board as the target, then you will have access to hardware settings. **Figure 14** shows one of the most important ones, pulse width modulation (PWM) timing. Here, you can adjust switching frequency, deadtime, and the interchannel delay in the switching waveform. These are parameters that will have a strong effect on distortion and efficiency—and they're the easiest ones to set poorly and transmute the GaN into carbon. So this is a powerful tool, which must be handled carefully.

With the hardware settings finalized, the developer can start working with the audio signal flow. **Figure 15** shows the default signal flow for the EVAL board, which the developer can modify. In this case, I added in a five-band equalization module to one channel by drag and drop using the software's GUI. **Figure 16** shows the particular EQ

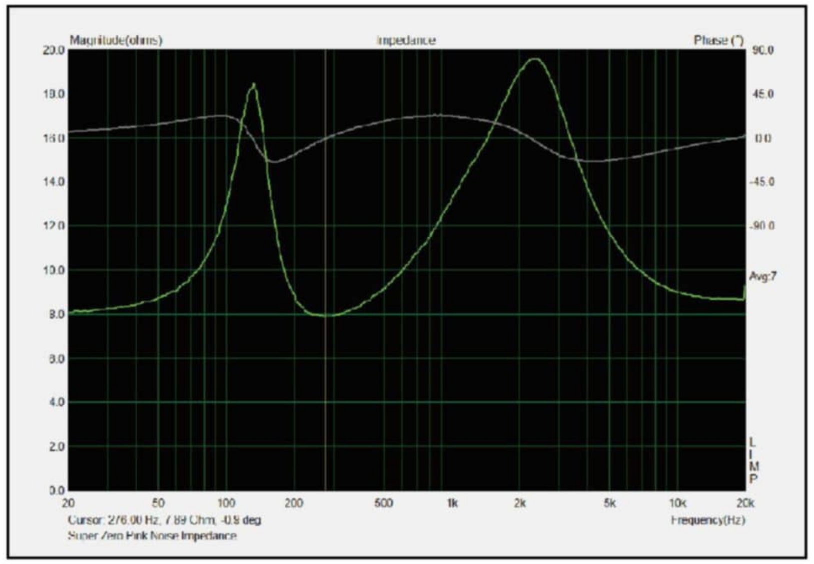


Figure 11: This is the impedance versus frequency of a loudspeaker used for a test load.

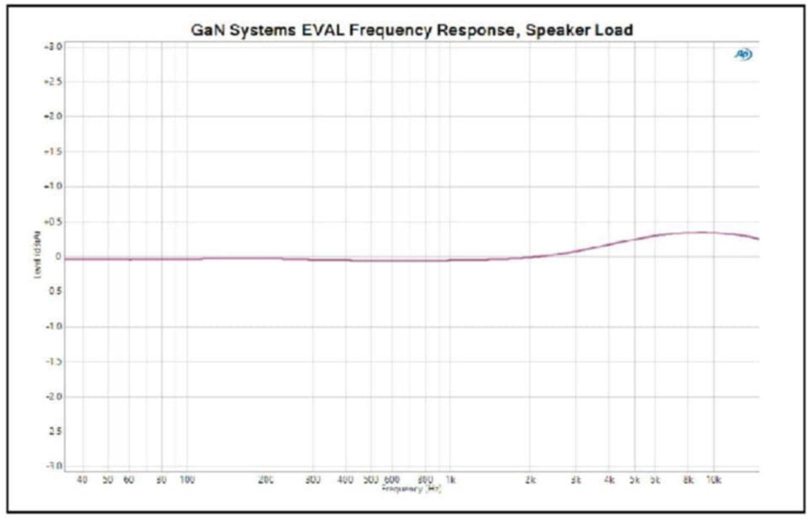


Figure 12: The EVAL frequency response into a speaker load shows a very minor deviation in the treble. Note the expanded scale to show this.

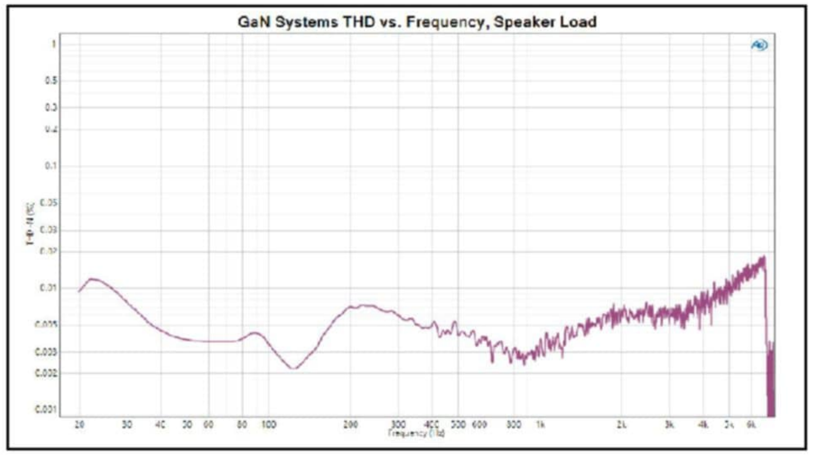
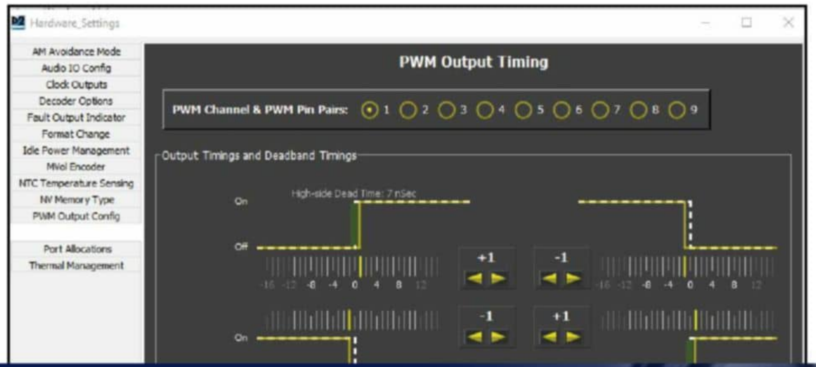


Figure 13: The variation of distortion with frequency into a speaker load shows that the EVAL board is perfectly comfortable driving a reactive load.



One note: The default signal flow is specific to the EVAL board. In the EVAL manual, EAS specifically warns about altering it (as opposed to adding functions). I, of course, knew better and proceeded to rework it the way I thought it should go. The dire warnings in the manual are more accurate than my self-assessment. Fortunately, Skip Taylor—my support contact indicated by GaN Systems—patiently walked me through reversing my errors and getting me back to a functional Square One. Moral: Believe the documentation.

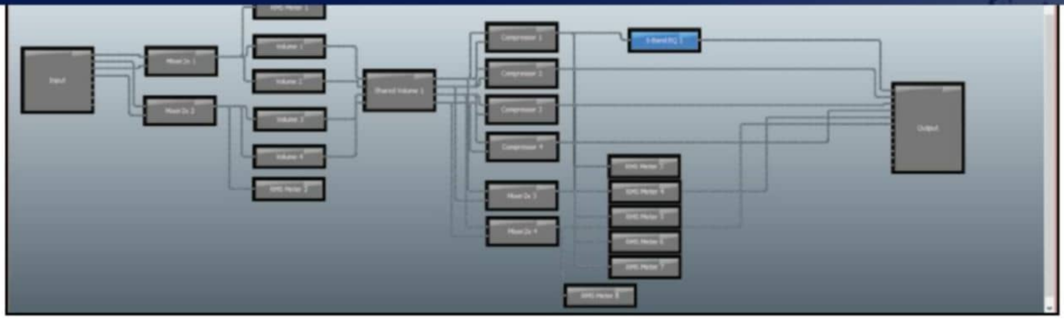


Figure 15: The default signal flow for the EVAL board can have functionality added in software- in this example, a 5 band EQ.

In theory, you could put this board in a box, run a few wires, and you'd have a fully featured high-performance integrated amplifier. Or use it to develop (or be built into) active loudspeakers using the DSP features. Or voiced instrument or PA amplifiers. The open-loop mode sounds remarkably like a good quality tube amp.

The intersection between excellent engineering and a remarkable device technology is not inexpensive ($1265), but delivers great performance, flexibility, and value for the price.

UPDATE: Prior to press time, GaN Systems announced the release of a Gen 2 version of this evaluation kit, which features a reduced footprint power supply and enhanced protection circuitry for the audio board. The company also claims a further increase in efficiency and thermal performance. GaN Systems, also debuted a new groundbreaking GaN-based 500W Class-D audio amplifier reference that merges GaN power transistors and Class-D controller technology from Dutch company Axign (implementing a digital control loop with feedback behind the output filter, across the loudspeaker terminals), enabling new interesting possibilities that *audioXpress* looks forward to evaluating.

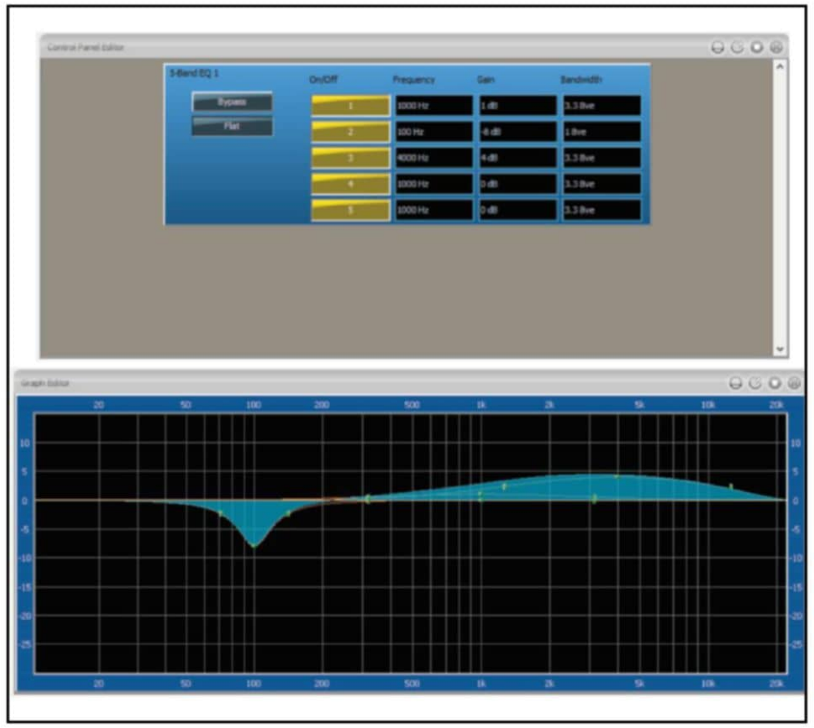
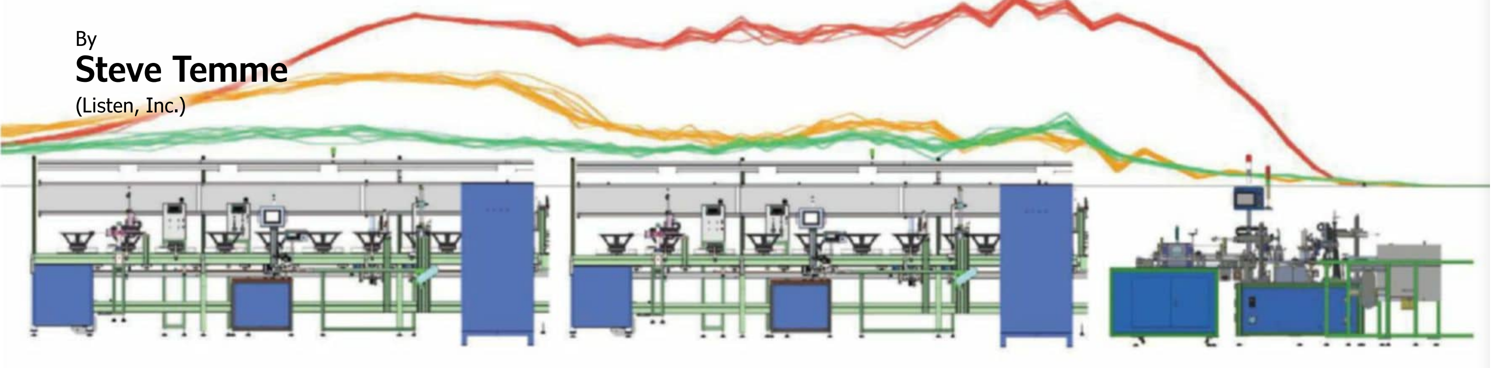


Figure 16: The EQ function can be controlled graphically or by entering values in a table.

The Evolution of Production Line Rub & Buzz Measurements

By
Steve Temme
(Listen, Inc.)



Rub & Buzz defects in loudspeakers, headphone drivers, and microspeakers generate an immediate perception of poor quality. This article follows the development of production line Rub & Buzz detection, starting with the conventional methods that, although developed 25 years ago, are still valuable and well-used today, and ending with the latest perceptual algorithms.

Production line detection of loudspeaker Rub & Buzz defects is important because the harsh sounds created by such faults instantly result in customer perception of poor quality. Such defects are almost always due to manufacturing errors, for example, the loudspeaker spider or surround may not be properly glued, the voice coil lead wires could be the wrong length, or the voice coil may be rubbing. If not immediately identified and corrected, these faults cause high volumes of waste and low yield, or worse still, customer returns.

Not all Rub & Buzz faults detected by conventional mechanisms are audible—the masking effects of the human ear and background noises can render some of them inaudible even though they are clearly present when the signal is analyzed. In the quest for higher yields, it's valid to question whether a speaker with inaudible faults needs to be rejected. This is particularly true for smaller and lower-margin products (e.g., microspeakers used in mobile phones, laptops, and some smart devices). Over the past decade, this desire for lower costs and higher yields has driven the development of accurate perceptual distortion algorithms that reliably identify whether a fault is audible, rather than simply whether it exists.

This article follows the development of production line Rub & Buzz detection, starting with the conventional methods that, although developed 25 years ago, are still valuable and well-used today, and ending with the latest in perceptual analyses. The trade-offs between each are discussed, and the reasons why both conventional (higher-order harmonic) and perceptual distortion measurements provide valuable data points for end-of-line tests are explained. It should be noted that these days, production line Rub & Buzz measurements are equally likely to be used on headphone drivers or microspeakers as they are conventional loudspeakers, and perceptual measurements are rapidly gaining traction in the quest for higher yields of such lower-cost products.

What Causes Rub & Buzz

Rub & Buzz defects are almost exclusively introduced in the manufacturing process. The term "Rub & Buzz" refers to the presence of higher-order harmonics, which create a harsh buzzing sound that often becomes particularly obvious at higher playback levels. Such sounds are created when something abruptly impedes the free movement of the speaker mechanism, for example, loose glue joints, vibrating voice coil lead wires hitting another loudspeaker part (e.g., the cone), or an incorrectly centered and rubbing voice coil.

Naturally a speaker with Rub & Buzz faults sounds bad and should be rejected, but identifying such speakers is just one benefit of measuring this important parameter. Often, Rub & Buzz is caused by improperly configured production line settings, or even a drift in parameters over time. It can therefore affect large volumes of product. Continual monitoring of Rub & Buzz enables trends to be noted and any line adjustments to be made in a timely manner, preventing waste.

Higher Order Harmonic Distortion (HOHD) Measurement—1995

Measurement of higher-order harmonics is often considered a "traditional" Rub & Buzz measurement. This technique, as the name suggests, relies on the analysis of higher-order harmonics—usually the 10th and above. These harmonics contribute most significantly to Rub & Buzz as their distance from the fundamental frequency makes them less masked and more easily detected by the ear. They also tend to fall in the frequency range most sensitive to the ear—around 1kHz to 5kHz. This is not to be confused with the measurement of total harmonic distortion (THD)—a completely different parameter that measures overall speaker harmonic distortion with much of the dominant information coming from the larger second and third harmonics that are excluded from Rub & Buzz detection. Although a valuable metric in itself, THD aligns more to the speaker design rather than manufacturing defects.

Prior to the development of Listen's higher-order harmonic distortion (HOHD) measurement technique, there were two methods used for measuring it. The "tracking high-pass filter" method, shown in **Figure 1**, uses a high-pass filter to remove the lower harmonics (usually up to 5th) and sums the remainder to give an overall metric. This method does not distinguish between periodic distortion (Rub & Buzz) and transient distortion (loose particles—see the Sidebar) lumping all the individual harmonics plus transient distortion together. This means that while it gives an overall indication of distortion level, it reveals nothing about the origin of the distortion, a major drawback in terms of understanding manufacturing problems. It is also susceptible to background noise, often resulting in false positives on the production line.

The second method, used in the popular (at the time) Brüel & Kjær 2012 analyzer, sequentially measures individual harmonics using a tracking filter that moves from one harmonic to the next. While this provides accurate and detailed information on each harmonic, it is not fast enough to use on automated production lines.

The launch of SoundCheck in 1995 brought the first production line Rub & Buzz algorithm that is fast, has good immunity to background noise, and provides individual harmonic detail. Using Listen's proprietary Harmonictrak algorithm, it measures each harmonic discretely and simultaneously, and offers the user complete flexibility in which harmonics are analyzed. The advanced filtering in this algorithm (**Figure 2**) excludes the noise between the harmonics, which means that in addition to its unprecedented analysis speed, it is also highly immune to background noise—an important characteristic for production line use.

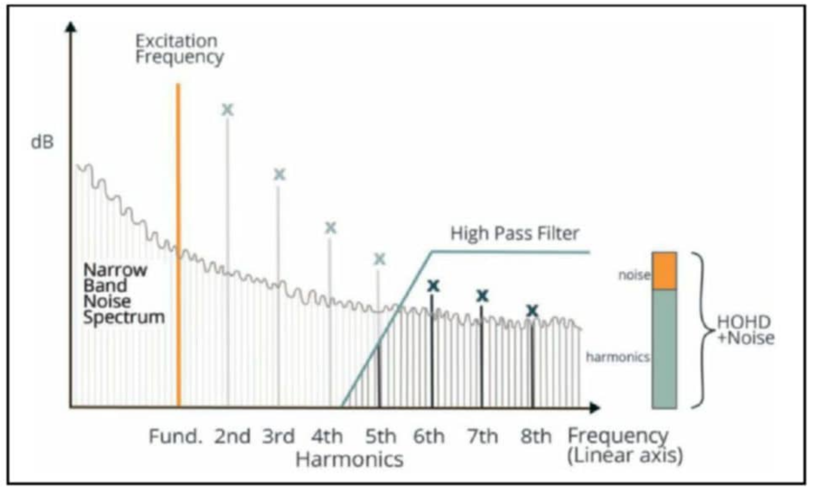


Figure 1: Rub & Buzz measurement using the high-pass filter method

The individual harmonic analysis offered by this technique reveals much about the failure mode, enabling a deeper analysis of the failure mechanism. High even-order harmonics are an indication of an asymmetrical nonlinearity, such as the voice coil not being centered in the magnetic gap. High odd-order harmonics are an indication of a symmetrical nonlinearity, such as clipping or reaching speaker cone and spider excursion limits.

High levels of harmonic distortion in the 10th to 15th order range usually indicate a rubbing voice coil. Distortion at the 50th harmonic and above is frequently caused by voice coil lead wires hitting the cone. Examining the harmonic response of speakers with known faults and carefully selecting the harmonics to analyze can reveal many discrete manufacturing defects. Trends can also be monitored over time, so that production line adjustments can be made if something is drifting out of specification over time. As with other production tests, it uses

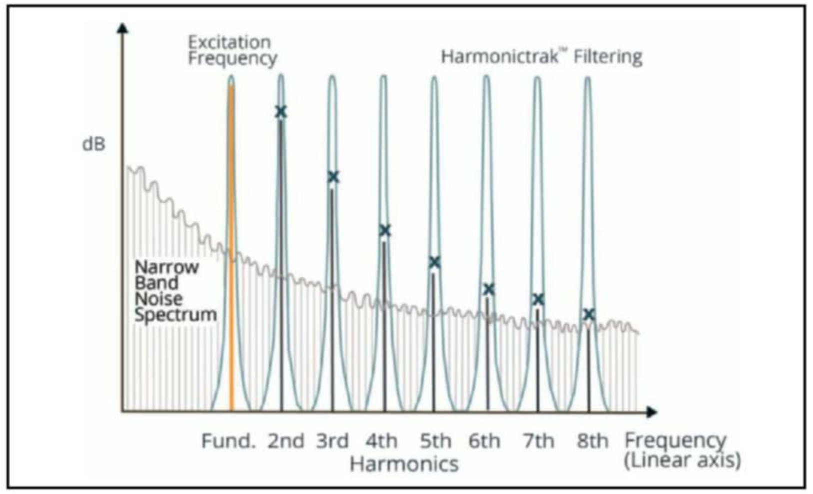


Figure 2: Demonstration of how the Harmonictrak algorithm eliminates the noise between harmonics to isolate the harmonics.

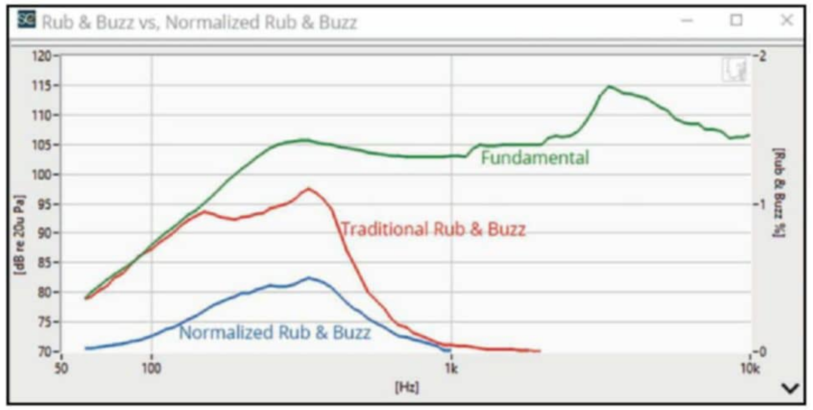


Figure 3: Comparison of traditional and normalized Rub & Buzz distortion

a stepped sine wave test signal, so the analysis is performed simultaneously with other end-of-line measurements.

Normalized Rub & Buzz Distortion—1996

The next innovation in Rub & Buzz measurement, Normalized Rub & Buzz distortion, came in 1996. I originally devised and proposed this method a few years earlier when working as an applications engineer at Brüel & Kjær, and it was the subject of a 1993 Audio Engineering Society (AES) paper, "How to Graph Distortion Measurements" [1]. SoundCheck was the first test system to include this as an out-of-the box software feature back in 1996, and the method was subsequently endorsed by Dr. Wolfgang Klippel in a 2004 *Journal of the Audio Engineering Society* (JAES) manuscript [2].

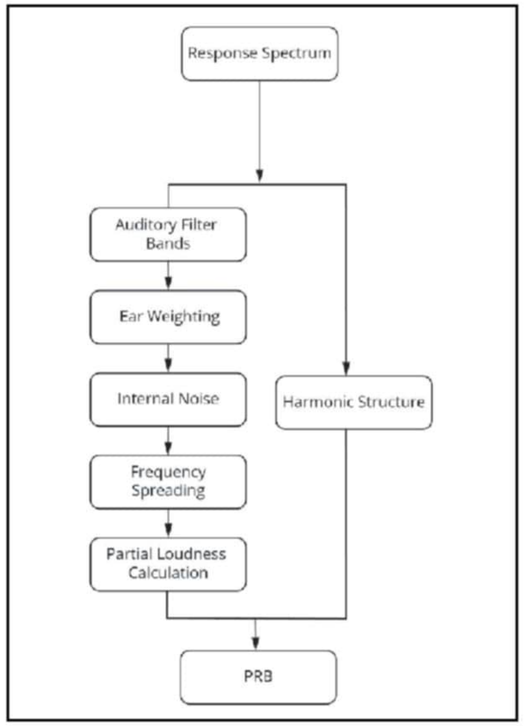


Figure 4: Schematic of original (2011) perceptual Rub & Buzz algorithm

This analysis is very similar to conventional Rub & Buzz distortion—in fact it is probably best described as an enhancement to it. It differs in that the harmonic levels are first compared to the fundamental at their measured frequency, then their ratio is plotted at their excitation frequency. This removes the effect of the frequency response from the distortion—a logical decision since the two are unrelated. This makes it easier to see the peaks in the distortion response, and easier to set limits as the frequency response of each speaker does not need to be accounted for. These are both significant advantages for production line use. Despite this technique being a quarter-century old, it is still used on production lines worldwide, and is just as valid today as when it was first conceived.

Figure 3 demonstrates why normalized Rub & Buzz is more revealing than conventional distortion measurement. The Fundamental (green line) shows a clear roll-off below 400Hz. This causes the conventional Rub & Buzz distortion (red line) to be higher relative to the fundamental in this range, which leads to an overestimation of the distortion level. When the effect of the frequency roll-off is removed, the shape of the curve (blue line) is different, with particular reduction in distortion at low frequencies. This makes the distortion measurement easier to interpret.

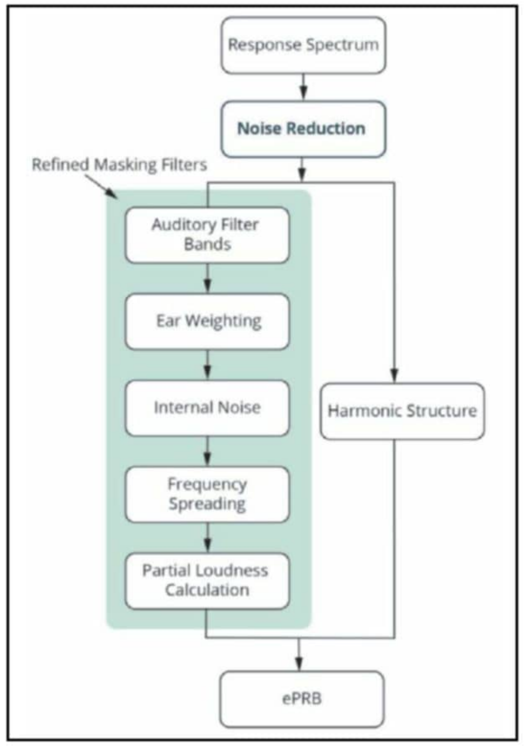
It is important to note that this Rub & Buzz detection method, as with conventional higher-order harmonic distortion, does not discern between audible and inaudible faults. This is advantageous in that it will pick up production line inconsistencies even if the fault is inaudible, making it a valuable diagnostic tool [3]. On the downside, speakers with inaudible faults may be unnecessarily rejected, thus reducing yield.

Perceptual Distortion—2011

By the late 2000s the quest for higher yields, particularly for the inexpensive, low-margin miniature speakers widely used in devices (e.g., mobile phones), spurred manufacturers to seek end-of-line tests that would only reject speakers with audible faults. This sparked the development of perceptual distortion algorithms [4]. In 2011, Listen launched the first perceptual Rub & Buzz distortion algorithm, called CLEAR.

This method (**Figure 4**) applies the psychoacoustic principles outlined in the ITU PEAQ standard for CODEC sound quality evaluation to create an algorithm that replicates human hearing. As with conventional Rub & Buzz measurement, it begins with a stepped sine wave stimulus. Two separate analyses are then carried out on the

Figure 5: Schematic of new (2022) Enhanced Perceptual Rub & Buzz (ePRB) algorithm



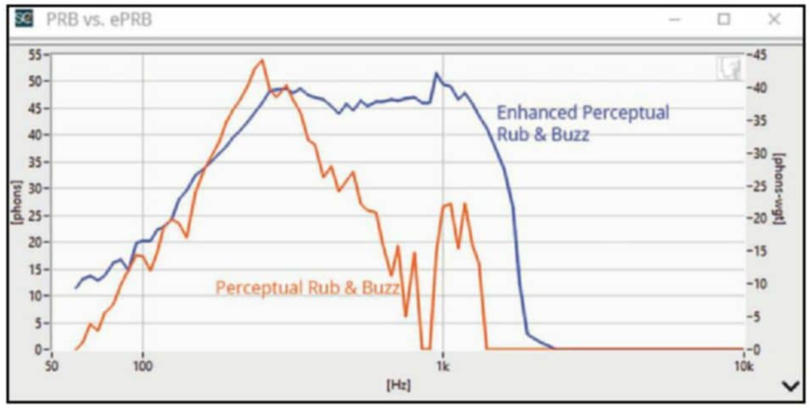


Figure 6: Comparison of original and new Perceptual Rub & Buzz algorithms

About the Author

Steve Temme is founder and President of Listen, Inc., manufacturer of the SoundCheck audio test system. Steve founded the company in 1995, and for the past 25 years the company has remained on the cutting edge of research in audio measurement, regularly introducing new measurement techniques, algorithms, and hardware to enable testing of all kinds of devices ranging from basic transducers to complex audio systems. In recent years, Listen's product development focus has been on creating the algorithms, interfaces and test methods to measure today's many and varied wireless and cloud-communicating devices. Prior to founding Listen, Steve worked for many years as an acoustic test and measurement applications engineer at Brüel & Kjær, and also as a loudspeaker design engineer at Apogee Acoustics. He holds a BSME from Tufts University, has authored numerous papers on acoustic testing, and has lectured extensively throughout the world.

response signal, which are combined into a numeric perceptual distortion measurement in Phons. In one analysis, the frequency response of the ear is modeled. First, auditory filter bands convert the fast Fourier transform (FFT) spectrum (constant bandwidth) to a Bark scale to replicate ear filtering. Next, an ear weighting filter compensates for the transfer function of the outer to inner ear, and finally, the internal noise of the ear (noise floor due to blood flow) is added. A frequency spreading function—a simplified mathematical representation of auditory masking curves—is then applied to mimic the psychoacoustic filters of the ear in hearing Rub & Buzz defects. These masking curves change with frequency and level. The fundamental and its masking effects are then removed to give the distortion of the speaker plus noise. This is summed over the frequency range to give the perceptual partial loudness for a single tone of the input signal.

On the other branch of the algorithm, the harmonic structure of the response is quantified using the power cepstrum (a cepstrum is a spectrum of a log spectrum). A strong and extended harmonic structure is a signature of Rub & Buzz.

In the final step of the algorithm, the result of the harmonic analysis is combined with the perceptual distortion to accentuate the Rub & Buzz, and return a harmonically weighted value for Rub & Buzz loudness in Phons [5, 6, 7].

This algorithm truly was a first in this industry, offering excellent correlation to human hearing. Although it was well received, its sensitivity to background noise made the repeatability inconsistent, and consequently made it challenging to set limits. This hampered its adoption in noisy production environments.

Overcoming these issues has not been an easy task; the fact that it was a good 10 years before any other test manufacturer launched a perceptual Rub & Buzz metric is testament to the magnitude of this challenge. Indeed, more than 10 years later, some of the other recently launched algorithms suffer from the exact same repeatability issues and dependencies as Listen's original CLEAR algorithm.

Enhanced Perceptual Rub & Buzz (ePRB) Distortion—2021

Enhanced Perceptual Rub & Buzz (ePRB) Distortion, Listen's newest perceptual Rub & Buzz algorithm has just been released with SoundCheck 20. This new algorithm, the product of several years of original research, builds on the sound psychoacoustic principles of the original CLEAR algorithm, but features two significant enhancements to improve performance (**Figure 5**). First, a new perceptual filtering algorithm significantly improves correlation to human hearing. Second, a novel noise-reduction algorithm offers significantly improved performance in noisy environments. This makes limit setting easier, and increases repeatability.

As with the original perceptual Rub & Buzz algorithm, this test method also uses a stepped sine sweep stimulus, enabling it to be measured simultaneously with other end-of-line quality metrics with no increase in test time. The underlying principles of this algorithm are well documented above and in several AES papers [5, 6, 7], so here the focus is the results of the recent enhancements.

Figure 6 compares the same speaker using both the original perceptual Rub & Buzz metric and the new enhanced version. The new algorithm shows much greater distortion in the mid-frequency range. This reflects the enhancements to the perceptual algorithm that make it more accurately model the behavior of the human ear in the highly sensitive mid-range frequencies. Hundreds of speakers were subjectively and objectively tested during the design of this algorithm, and the correlation with subjective opinion was significantly higher than with the prior version.

Next, we compare good, bad, and borderline speakers using Listen's original 2011 algorithm, Listen' new 2022 algorithm, and another recently launched perceptual Rub & Buzz algorithm. Three speakers of the same type—a typical small loudspeaker—were analyzed.

One speaker is good, one has high levels of Rub & Buzz distortion, and the third exhibits some distortion, but it is not as audible as in the speaker with high distortion. These three particular devices were chosen from a much larger batch as examples of good, mildly distorted, and severely distorted output in order to clearly show contrast in the results. The three speakers were subjectively assessed by a panel of engineers experienced with listening to distortion artifacts, and the results are shown in **Table 1**.

Figure 7 shows the results of these three speakers tested 10 times each in an ordinary room, and measured with the three different algorithms outlined in Figure 7— Listen's original Perceptual Rub & Buzz (a), the new enhanced Perceptual Rub & Buzz (b), and the other recently launched algorithm "Other Perceptual Algorithm" (c).

In Figure 7a, the original (2011) perceptual algorithm visually indicates the distortion faults, clearly showing correlation to subjective listening results. The bad speaker shows the highest level of audible distortion, the good speaker has minimal measured distortion, and the borderline speaker is somewhere in the middle. However, when the speakers were measured 10 times, small variations in the results are observed. This is primarily due to background noise, and in a production environment where there is only one test sweep on each device, this variation makes limit-setting challenging.

Figure 7b shows Listen's new (2022) ePRB algorithm, also with the exact same 10 measurements for each speaker. It is immediately apparent that the results are more consistent, as is expected with less sensitivity to noise. This improved repeatability makes it considerably more reliable than the original algorithm and easier to set limits, therefore making it more suitable for production line use. Additionally, the enhancements to the algorithm reveal increased detail in the ear's highly sensitive 500Hz to 2kHz range, which more accurately mirrors human perception.

Figure 7c shows another recently launched perceptual Rub & Buzz algorithm. This demonstrates similar repeatability issues to SoundCheck's 2011 algorithm, and, interestingly, it measures greater audible distortion at low frequencies for the speaker

Speaker	Determination	Listener Comments
97	Good	No audible distortion
76	Borderline	Some audible Rub & Buzz at low frequencies only
57	Bad	Very audible Rub & Buzz over most of the frequency range

Table 1: Subjective analysis of measured loudspeakers

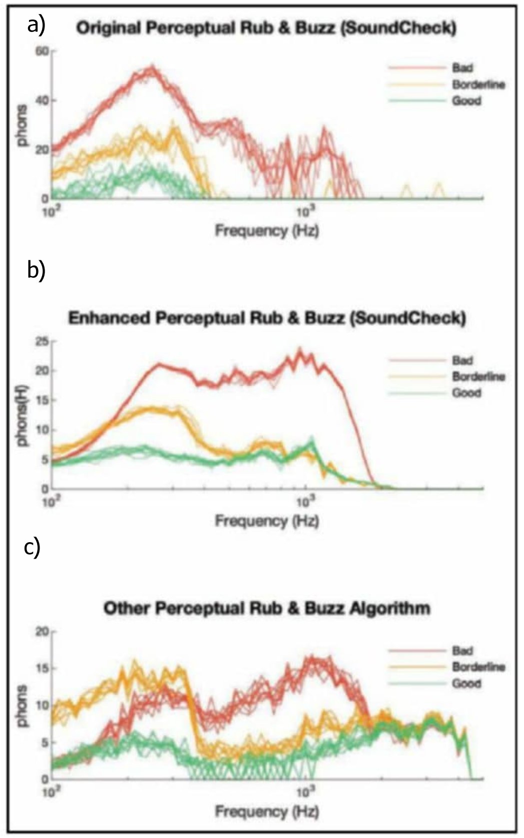


Figure 7: Comparison of three different perceptual Rub & Buzz algorithms.

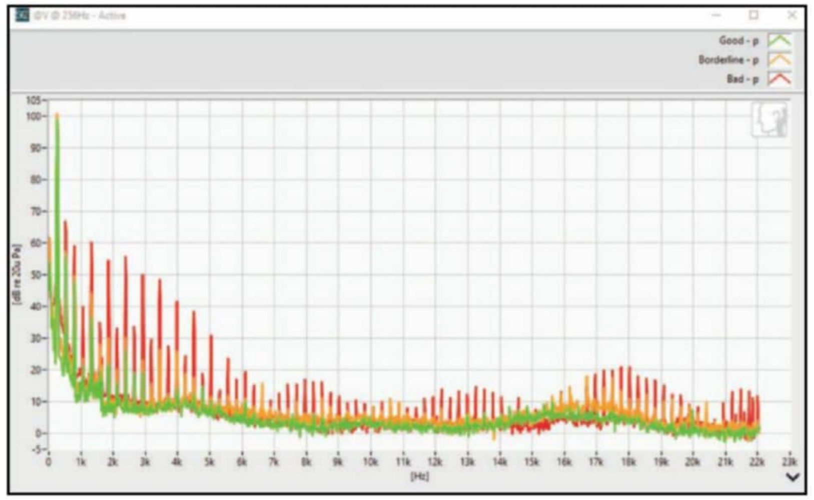


Figure 8: Conventional (HOHD analysis) of three test speakers

that was subjectively assessed as borderline than the one subjectively assessed as bad.

These results are validated by examining the harmonic distortion using "conventional" FFT analysis (**Figure 8**), as well as correlating them to listener experience. To simplify both the listening experience (it's very hard for even a trained listener to hear distortion in a fast sweep) and the harmonic spectrum, a fixed frequency of 256Hz was used to evaluate the same three speakers. This frequency was chosen as it reflected the frequency at which both bad speakers exhibit a peak in perceptual distortion measurement and therefore would be the easiest to hear.

The distortion spectra clearly shows minimal high-order harmonic distortion for the good loudspeaker (green line), which correlates well with the perceptual measurement, as one would expect. Both the bad and borderline speaker show elevated high-order harmonic distortion, and again, in line with Listen's perceptual algorithms, the bad one (red) clearly has more distortion than the borderline one throughout the spectrum. At no point does the borderline (orange) speaker exhibit more distortion. As one might expect, the perceptual distortion algorithm amplifies the relative magnitude distortion in the highly sensitive part of the ear's range.

It is interesting that the "other" perceptual algorithm shows little difference in the maximum distortion audibility of the two speakers that exhibit Rub & Buzz, although human listeners perceived a significant difference in distortion level. Furthermore, at the 256Hz frequency measured above, it actually showed greater distortion with the borderline speaker (orange line) than the bad speaker (red line), which does not correlate with either the HOHD measurement or the listener experience.

Distortion Parameters for Assessing Production Line Defects

Generally speaking, there are three valuable distortion parameters that should form part of your end-of-line test for assessment of manufacturing

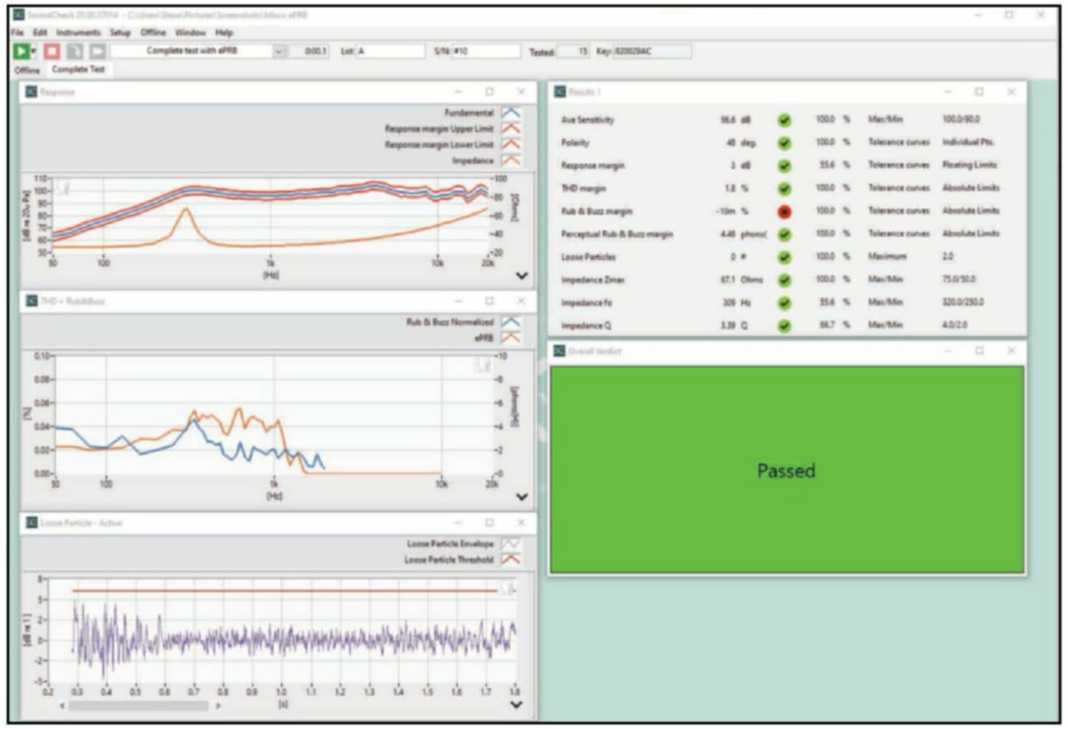


Figure 9: Complete end-of-line loudspeaker test, demonstrating how a speaker might pass Perceptual Rub & Buzz metrics and the overall test while failing conventional Rub & Buzz because it is not audible.

defects: individual harmonics distortion, loose particles, and perceptual Rub & Buzz. Each serves a particular purpose.

Normalized Rub & Buzz (or HOHD) will identify any fault on your production line that is leading to a rubbing or buzzing device, whether you can hear it or not. Trends can be monitored over time to catch any production line drift before speakers start being rejected. An analysis of the specific harmonic that is causing the device to fail can provide valuable insight into what the specific problem is, accelerating production line troubleshooting and enabling problem areas to be addressed. Although standard discrete higher-order distortion measurement will also suffice, the normalized version is more accurate and more revealing. It also makes it considerably easier to see the peaks in response and set limits—important characteristics for production line use.

Loose particle detection rejects speakers with transient faults. This tends to be caused by the ingress of foreign particles into the speaker during manufacture and tends to occur on a more random basis. Nevertheless it still sounds unpleasant to the human ear and speakers with this defect in the audible range should always be rejected.

Perceptual Rub & Buzz is an indicator of whether the consumer will be able to hear the fault. In some situations, particularly with lower cost devices, a knowledge of whether the distortion is audible may affect the decision on whether to pass or fail it. In other words, if a speaker's Rub & Buzz defect is not audible, in many cases it may be acceptable to ship it, since the consumer will never hear it. This reduces manufacturing waste, and therefore increases yield/profitability for the line.

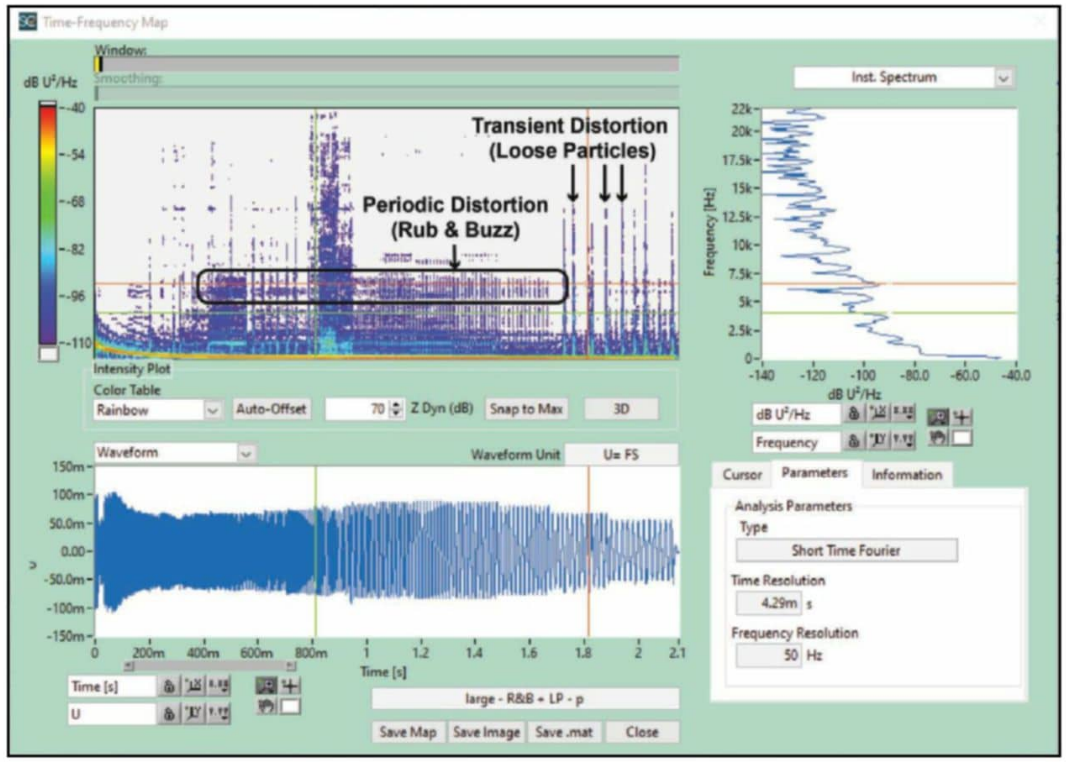


Figure 10: Transient distortion (loose particles) shows up in the time domain, whereas periodic distortion (Rub & Buzz) shows up in the frequency domain.

Loose Particles

It is important that Rub and Buzz measurements, perceptual or otherwise, are not confused with Loose Particles or other randomly occurring transient defects. Like Rub & Buzz, loose particles cause an unpleasant clicking or rattling sound within the speaker, and affected speakers should be similarly rejected. However the failure modes that cause loose particle defects, and the method used to detect them, are completely different. Loose particles are caused when foreign particles such as glue or magnet fragments become trapped in the gap behind the diaphragm or dust cap during the manufacturing process. The random clicking, popping or rattling noises made by these trapped particles cause transient, rather than periodic distortion. This is identified by time-envelope analysis of the waveform response in the time domain (**Figure 10**). Detection thresholds are set for the magnitude, duration, and number of transients to quantify their level and number over time. Limits and filters are applied to separate device transients from background noise. Fine tuning with these tools customizes the algorithm for your product and manufacturing environment [8, 9].

All three measurements use a stepped sine stimulus. This means that they can be measured and analyzed simultaneously (along with other production line metrics such as frequency response, THD, polarity, impedance including Thiele-Small parameters and more) in a single sweep (**Figure 9**). Typically this takes as little as one second, making measurement of all of these parameters fast and easy on today's high-speed production lines. Whatever other end-of-line measurements are made, these three will enable you to accurately identify common manufacturing defects and remove the affected speakers.

The Future of Production Line Distortion Measurements

Perceptual measurements are the "holy grail" of production electroacoustic measurements because of their potential to increase the yield of high-quality products in a cost-driven marketplace. This will likely result in many innovations and algorithms in the next few years that better correlate manufacturing defects to human auditory perception, and will likely include some overall psychoacoustic and distortion metrics.

References

[1] S. Temme, "How to Graph Distortion Measurements," presented at the Audio Engineering Society (AES) 94th Convention, March 1993.

[2] W. Klippel, "Measurement and Amplification of Equivalent Input Distortion," *Journal of the Audio Engineering Society* (JAES), Volume 52, Issue 9, pp. 931–947, September 2004.

[3] D. Foley, Dr. R. Celmer, B. Sachwald, J. Anthony, T. Pagliaro, and S. Thompson, "Higher Order Harmonic Signature Analysis for Loudspeaker Defect Detection," presented at the 117th Audio Engineering Society (AES) Convention, October 2004.

[4] S. Temme, P. Brunet, and D.B. Keele Jr., "Practical Measurement of Loudspeaker Distortion Using a Simplified Auditory Perceptual Model," Presented at the 127th Audio Engineering Society (AES) Convention, October 2009.

[5] S. Temme, P. Brunet, and B. Fallon, "Practical Implementation of Perceptual Rub & Buzz Distortion and Experimental Results" Presented at the 129th Audio Engineering Society (AES) Convention, November 2010.

[6] S. Temme, P. Brunet, and P. Qarabaqi, "Measurement of Harmonic Distortion Audibility Using a Simplified Psychoacoustic Model," presented at the 133rd Audio Engineering Society (AES) Convention, October 2012.

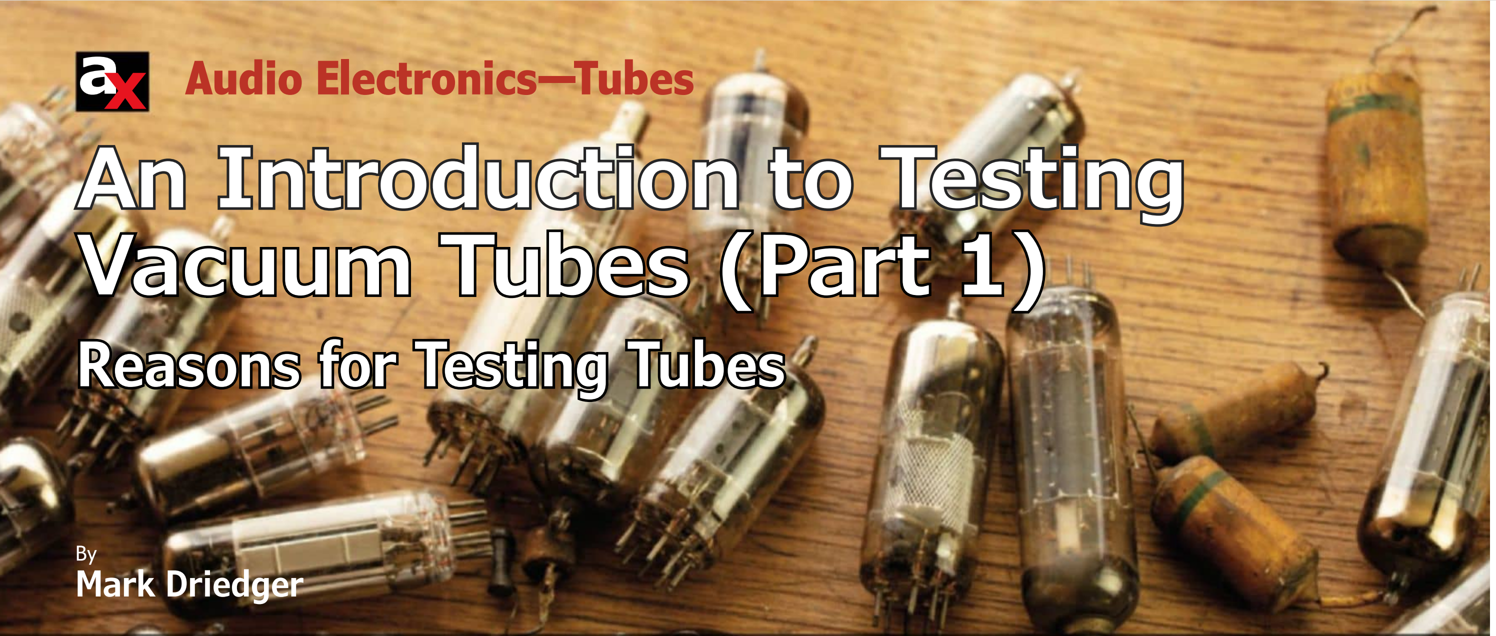
[7] S. Temme & P. Brunet, "Measurement of Harmonic Distortion Audibility Using a Simplified Psychoacoustic Model – Updated," presented at the 51st Audio Engineering Society (AES) International Conference, August 2013.

[8] P. Brunet, E. Chakroff, and S. Temme, "Loose Particle Detection in Loudspeakers," presented at the 115th Audio Engineering Society (AES) Convention, October 2003.

[9] P. Brunet and S. Temme, "Enhancements for Loose Particle Detection in Loudspeakers," presented at the 116th Audio Engineering Society (AES) Convention, May 2004.

Resource

Listen, Inc., www.listeninc.com



An Introduction to Testing Vacuum Tubes (Part 1)

Reasons for Testing Tubes

By
Mark Driedger

If you are a fan of tube audio, sooner or later you will need to test tubes. In this two-part article, Mark Driedger examines why and how we test tubes. He refers to several commercial and DIY testers; however, this is not intended to be an exhaustive survey. Please be aware that most tube testers expose the user to hazardous voltages and you should exercise extreme caution. If the tester has a plug board, do not touch it while the tester is powered. The unused tube sockets will have high voltage on them, as may the plate/grid top cap, which is likely dangling from a frayed and cracked wire on the 70-year-old tester you just bought on eBay. You are responsible for your safety—not the author or *audioXpress*.

Reasons to Test Tubes

I am going to assume that you know the basics of how tubes work and will focus the discussion on audio amplifier applications. Tube testers of yore had to deal with a bewildering array of tubes and circuits, especially for TV, radio, and military use. Fortunately, the situation is simpler for modern tube audio amplifiers. Tube rectifiers are not covered, though some of the same principles apply.

Before I get your hopes up too high, there is no ideal tube test or tube tester. The "goodness" of a tube is very tightly connected to your application. A tube can work fine in one circuit and perform poorly in another. In fact, a good circuit is one that is insensitive to tube variation. As an example, an amplifier using self-bias (bias resistor in the cathode leg) or a constant current source to set bias is inherently less sensitive to tube variation and aging than one using fixed bias or grid leak bias.

In Robert Tomer's classic text *How to Get the Most out of Vacuum Tubes*, his advice is to "… never test any tube if you can possibly avoid doing so" (page 153).

Plate current	Reduced from initial value to 20mA
Mutual conductance	Reduced from initial value to 24.5mA/V
Negative grid current	Increased from initial value to >1µA

Table 1: A Telefunken datasheet provides end-of-life test parameters for the Telefunken D3A.

Periodic testing usually decreases tube life by stressing the mechanical integrity of the glass bulb during insertion and removal. Some tests are inherently damaging. That said, when a circuit is not performing well it is necessary to determine if the tube is the cause.

There are four main reasons for testing tubes: to determine if a tube is "faulty," to determine if a tube is "worn out," to characterize a tube's parameters, and to match multiple tubes to each other.

Tube Faults: A tube can be faulty due to obvious physical damage to the glass envelope or pins, an open filament, shorts between elements, and excessive gas. An open filament is much like a burnt-out light bulb and happens for many of the same reasons. Filament life is shortened by excessive filament voltage, the thermal stress of repeated power cycling, mechanical shock, and loss of vacuum. Open filament = cold tube = no glow, so you do not need a tube tester to figure this one out. Shorts are caused by small pieces of debris left in the tube bulb during manufacturing, contamination/failure of the mica insulation, or from arcing induced by excessive voltage between closely spaced tube elements. The "shorts" are not necessarily low resistance or continuous. For example, the Heathkit IT-17 classifies anything less than 250 KΩ as a short. Residual levels of gas exist from manufacturing, and more gas is created during operation when the metals in the tube are bombarded by positive ions (see page 26 of Tomer's book for a detailed discussion). Manufacturers use a metallic flashing called a getter (usually barium or magnesium) on the inside of the tube bulb to absorb gas. A white getter is a sign the vacuum has been lost and the tube is faulty. Gas may cause a blue or pink glow within the tube, especially at high temperatures and voltages.

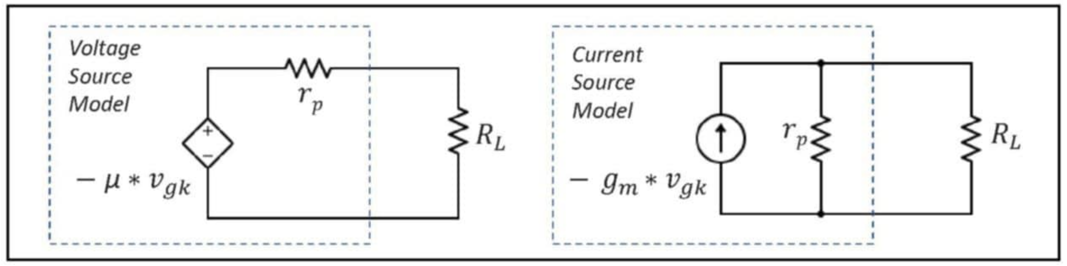


Figure 1: Small signal tube models are shown using the AC tube parameters.

Wear Out: The main wear-out mechanism is a loss of cathode emission over time. Cathodes are coated with some combination of materials (e.g., barium oxide, strontium oxide, and calcium oxide), which are good emitters of electrons when heated. The cathode creates a space charge of electrons, from which the plate and the screen grid currents are derived. If the cathode does not emit enough electrons to sustain the space charge, then electrons are stripped directly from the cathode and damage its surface. The damaged areas emit fewer electrons and exacerbate the problem.

A similar mechanism called cathode striping results when plate current is drawn before the heater comes to full operating temperature. Cathode poisoning is a third wear out mechanism. If the tube is heated for an extended period with no plate current, then an interface layer of oxide forms between the nickel cathode and the emissive material deposited on it. The interface layer reduces emissivity and increases noise. Special materials were used to reduce this effect, especially in tubes meant for computers, which were heated continuously to avoid thermal cycling induced heater failures.

Increasing grid leakage current over time is a secondary wear-out mechanism. Material boiled off the emitter can collect on the grid, causing it to operate like a smaller secondary cathode and share the plate current (grid emission). The negative grid will also attract positive ions, causing grid current. Both effects increase with temperature. Tube datasheets specify a maximum grid circuit impedance (the impedance seen by the grid looking into the external circuit). Grid leakage current causes a voltage drop across this impedance, which shifts the bias point of the circuit. As the grid impedance increases, this effect becomes more pronounced, and the circuit becomes more sensitive to grid leakage. In very low-cost equipment (e.g., five-tube AM radios), it was common to establish bias solely by the grid leakage current flowing through a high value grid resistor (grid leak bias). What was gained in cost was lost in reliability, and this approach is not used in modern amplifiers.

Interestingly, very few tube datasheets provide any guidance on end-of-life testing. Telefunken is an exception as this example (**Table 1**) from its D3A datasheet shows.

Characterization: Tubes are characterized for their DC and AC (small signal) behavior. The DC parameters determine the operating point (bias), and the AC ones predict how it will behave as an amplifier. They are useful in designing circuits, matching, and in detecting wear out.

DC behavior is measured at one operating point by fixing any two of grid-cathode voltage, plate-cathode voltage and plate current, and then measuring the third. Curve tracers (discussed in Part 2 of this article) show DC behavior across a range of operating points.

There are three interrelated AC parameters: transconductance (g_m), amplification factor (µ), and plate resistance (r_p). Transconductance is the ability of grid voltage to cause plate current changes (at fixed plate voltage). The amplification factor is the ability of grid voltage to cause plate voltage changes (at fixed plate current), and plate resistance is the ability of plate current to cause plate voltage changes (at fixed grid voltage). Stated mathematically:

$$g_m = -\left.\frac{\Delta i_p}{\Delta v_g}\right|_{V_p\ constant}$$

$$\mu = -\left.\frac{\Delta v_p}{\Delta v_g}\right|_{I_p\ constant}$$

$$r_p = \left.\frac{\Delta v_p}{\Delta i_p}\right|_{V_g\ constant}$$

The three are inter-related by $\mu = g_m \times r_p$ and a simple AC small signal model of a tube can be defined either as a voltage dependent voltage source (using µ), or a voltage dependent current source using g_m. In both models, R_L is the equivalent load resistance in the plate circuit (external to the tube). The two forms are Thévenin/Norton equivalents of each other (**Figure 1**).

The AC parameters are "small signal" ones, defined mathematically as partial derivatives and generally measured with small AC signals or small step changes. All three vary dramatically with the

About the Author

Mark Driedger has been experimenting with tube audio since 1980. He completed an MSc in electrical engineering at the University of Waterloo, Canada, and has worked in the telecom industry for more than 30 years in technical and executive roles. His other passions include woodworking, the guitar, and spending time with his wife, children, and grandchildren.

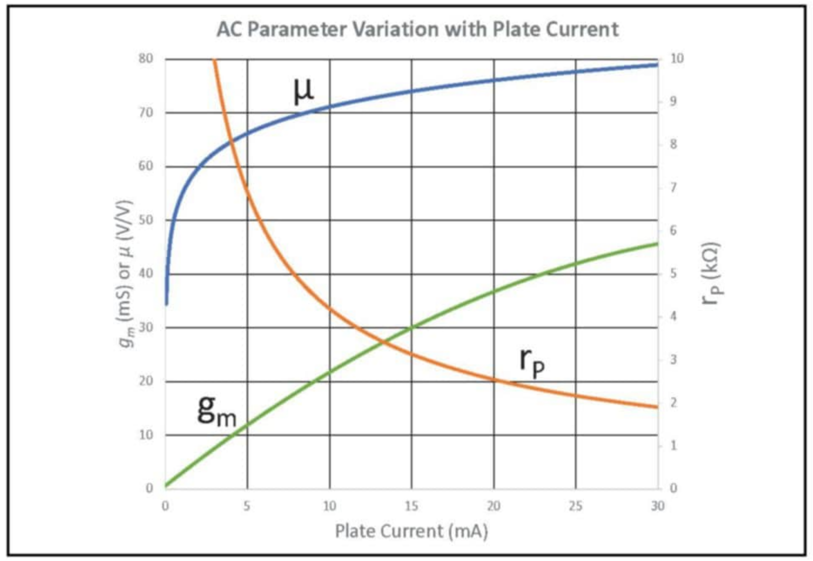


Figure 2: This graph shows an AC parameter variation with plate current.

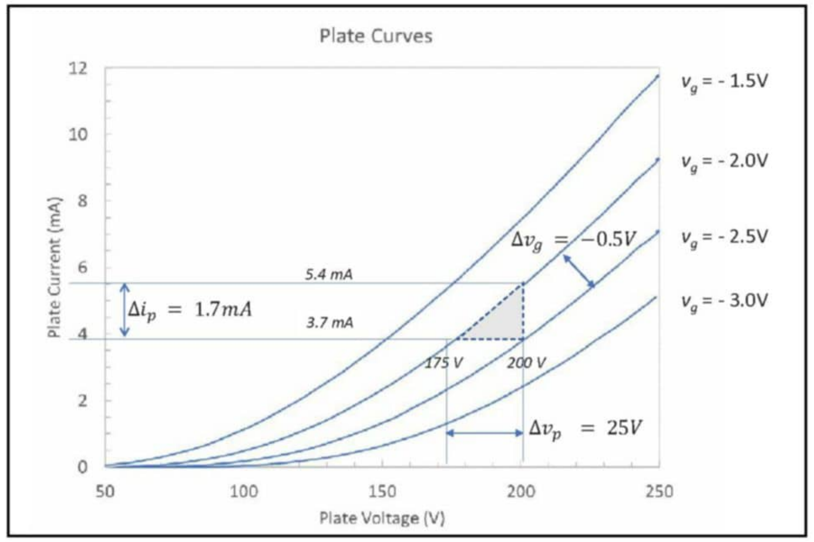


Figure 3: Typical triode plate curves are shown with a graphical estimate of tube parameters.

Resources

J. Didden, "The Internal Life of Vacuum Tubes, *audioXpress*, September 2020, https://audioxpress.com/article/the-internal-life-of-vacuum-tubes

J. Didden, "Book Review: Getting the Most Out of Vacuum Tubes," August 2020, https://audioxpress.com/article/book-review-getting-the-most-out-of-vacuum-tubes

A. Preisman, *Graphical Constructions for Vacuum Tube Circuits*, McGraw Hill, 1943.

Telefunken D3a datasheet, https://frank.pocnet.net/sheets/128/d/D3a.pdf

R. Tomer, *Getting the Most out of Vacuum Tubes*, Howard W. Sams, 1960, reprinted by KCK Media Corp., 2019, https://cc-webshop.com/products/getting-the-most-out-of-vacuum-tubes

DC operating point and, to be useful, should be measured at the intended plate current. **Figure 2** shows typical variation from which it is obvious that specifying g_m, for example, without nailing down the plate current is futile. In pentodes, μ is large and not a useful measure.

As a tube ages, g_m will tend to drop and r_p will tend to increase in roughly the same proportion, resulting in relatively constant μ. Near the end-of-life g_m drops more rapidly and μ decreases, leading to less circuit gain.

Another way to express the DC and AC behavior of the tube is with a graph containing multiple "plate curves" of plate current vs. plate voltage, at various grid-cathode voltages. From these curves, the DC operating point can be derived, either by reading data points or graphically for a given type of circuit. Before computer modelling and spreadsheets, most tube circuit design was done graphically from the plate curves. Albert Priesman dedicates an entire book to graphical design of circuits. The basics are simple and there are many tools on the web for Class A amplifier design from tube curve data. The AC parameters are estimated as slopes between adjacent grid voltage lines. **Figure 3** shows sample plate curves, and the AC parameters, which are estimated graphically from them.

$$g_m = \left.\frac{\Delta i_p = 1.7mA}{\Delta v_g = -0.5V}\right|_{V_p = 200V} = 3.4mS$$

$$r_p = \left.\frac{\Delta v_p = 25V}{\Delta i_p = 1.7mA}\right|_{V_g = -2V} = 14.7k$$

$$\mu = -\left.\frac{\Delta v_p = 25V}{\Delta v_g = -0.5V}\right|_{I_p = 3.7mA} = 50\,{}^{V}\!/_{V}$$

$$\mu = g_m \times r_p = 3.4mS \times 14.7k = 50\,{}^{V}\!/_{V}$$

There is a second form, called transfer curves, which are plots of plate current (vertical axis) vs. grid voltage (horizontal axis), with each curve at a different plate voltage. The slope of these curves is g_m and these curves contain the same information as the plate curves, just in a different representation (**Figure 4**).

The datasheet will list a variety of other parameters; however, few of them are routinely tested. Noise can be a concern in high gain circuits (e.g., microphone and phono preamplifiers). It is caused by the random absorption of positive ions by the grid among other mechanisms. Tubes are also sensitive to shock and vibration, a mechanism called microphonics. Mechanical forces will distort the relative spacing of the tube elements and affect the plate current and this can cause a sustained ringing or oscillation. The military spent many of our tax dollars looking for ways to reduce microphonics in tubes used in rugged environments (e.g., aircraft and missiles).

Matching: Push-pull circuits perform better if the two tubes in them are matched, meaning they have similar characteristics. Matched devices produce better balance and lower distortion. In a stereo amplifier, the two channels should also be matched to achieve similar performance. Matching seems simple enough. However, there are many parameters that could be matched, and they all vary with the DC operating point. Tubes can be matched for DC operating point (grid voltage at a specified plate current and voltage, for example), and/or AC characteristics (typically g_m). Tube curves can also be compared qualitatively or using some quantitative measure of their similarity.

Summary

At this point I hope you have a better understanding of how tubes fail and what we test for. In the next issue, I will dive into how to test tubes and the different types of testers that are available.

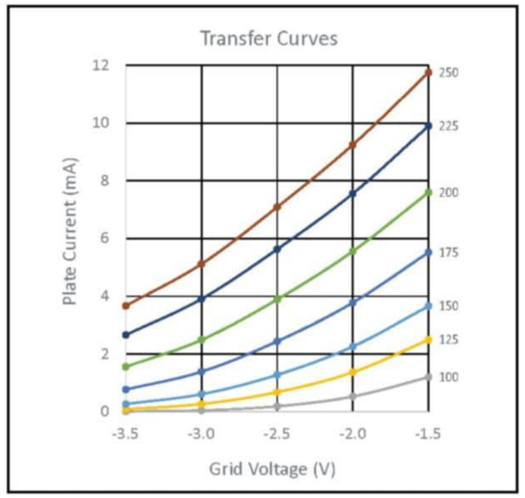


Figure 4: Here is a representation of typical transfer curves.

Going Further

Robert Tomer's book is a must-read. He covers tube lifetime and failure in more detail and provides advice on how to maximize tube life. This classic text is available from audioXpress as a reprint, https://cc-webshop.com/products/getting-the-most-out-of-vacuum-tubes, and Jan Didden provides a good review called "The Internal Life of Vacuum Tubes," which was originally printed in *audioXpress*, August 2020. Didden has a related article on tube lifetime, which is also recommended.

This article is the first chapter of a fascinating DIY journey that Ethan Winer takes through the world of synthesizers. More specifically, a guitar-controlled, all-analog synthesizer project that includes a "time machine" sample and hold section, input signal detectors, a noise gate, oscillators, envelope generators, filters, signal converters, and more. Readers are welcome to try any of the modules or build the complete project.

Building a Guitar-Controlled Synthesizer

The Sample & Hold Time Machine

By
Ethan Winer

Photo 1: Ethan built this synthesizer in 1970. The power supply heatsinks on top were originally inside the case. But the heat they generated caused the synthesizer oscillators to drift out of tune. (Photo credit: Kevin Byron)

I've always been fascinated with synthesizers. I grew up in the 1960s when the earliest commercial models were analog, starting with the Moog Modular, a complex expensive beast. Then later the simpler and less expensive MiniMoog and ARP 2600 models were introduced, soon followed by more models from other companies. As a poor teenager I couldn't afford to buy those, but I had a keen interest in electronics. So I set out to build my own. By the 1970s, I had built several synthesizers starting with a basic version controlled by a Hewlett-Packard (HP) computer. I never had formal schooling for electronics, but working as a technician at various companies gave me access to real engineers who were kind enough to answer my endless beginner questions.

At one company, I became friends with Leo Taylor, who later got a fabulous job as a service tech for HP. Leo's workbench included all the latest HP test equipment—storage oscilloscopes, signal generators, distortion analyzers, and more. Over a period of several years we met at his company every Monday night to design synthesizers and other audio devices.

Synthesizer Basics

One of the basic principles of analog synthesizers is voltage control. Rather than use only knobs to control the volume, or the pitch of an oscillator, or the cut-off frequency of a filter, in a synthesizer these are also varied by control voltages. So when you press a key on the keyboard, a voltage corresponding to that note is sent to a voltage-controlled oscillator (VCO). Another voltage that ramps up and down at user-defined rates feeds a voltage-controlled amplifier (VCA), and yet another varying voltage might go to a voltage-controlled filter (VCF) to add animation to the sound.

In what was surely one of the earliest attempts to sync a computer with an analog tape recorder, in 1970 Leo and I recorded a short tune, a modern three-part invention written by a friend of mine. We started by recording 60Hz hum onto one track of my Ampex AG-440 4-track professional recorder. That hum was then squared up—overdriven to clip and become a square wave—and played from the recorder into the control port of a small HP computer Leo borrowed from work. In 1970 a "small" computer was the size of a dormitory refrigerator!

Leo wrote a program in binary machine code to read the 60Hz sync tone recorded on Track 1, entering it one instruction at a time on the computer's 16 front-panel toggle switches. There were no diskettes or portable hard drives back then either. The musical note data for the tune was also entered via the computer's toggle switches. It took us 11½ hours to program the computer, enter the song data, and record the three tracks one at a time from control voltages output by the computer. You can hear the result in the file 1970_synth.mp3 listed in Resources.

In 1974, I built another synthesizer, shown in **Photo 1** and **Photo 2**. This one was much more ambitious than the first and included a 61-note keyboard, four audio oscillators, two VCFs and two VCAs, a low-frequency oscillator (LFO) for vibrato and tremolo, and portamento (glide from note to note). A 10-step sequencer was also included, along with circuitry to split the keyboard at any point so both synthesizer "halves" could be played simultaneously for a whopping two-note polyphony.

It took Leo and me two years to design this synthesizer, and I spent another nine months building it in the evenings. Leo was the real brains, though by the end of this project I had learned enough to design all the keyboard and control circuits myself. Fourteen separate plug-in circuit cards hold the various modules, and each card was hand-wired manually using crimped wiring posts and 24-gauge buss wire with Teflon sleeving.

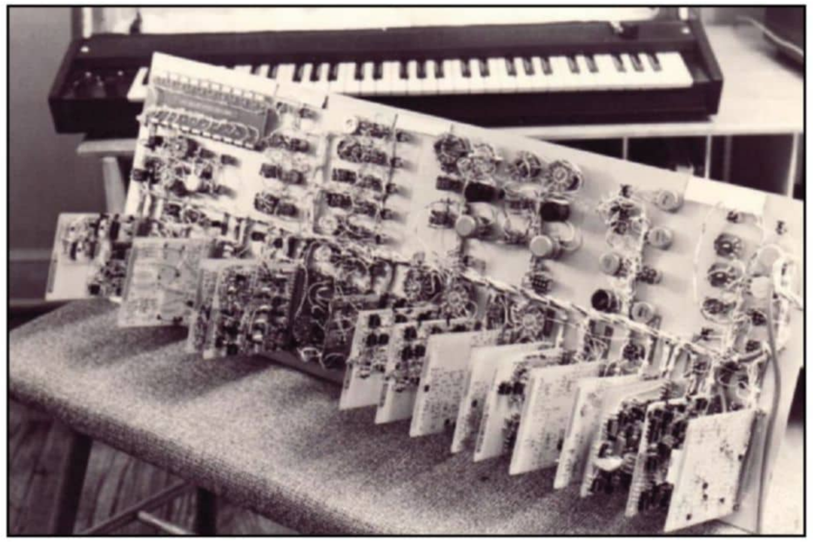


Photo 2: It took nine months working every evening to build this monster. (Photo credit: Kevin Byron)

Designing a Guitar Synthesizer

Fast forward to early 2021. I'm not much of a keyboard player, though I can pound out simple melodies and chords at half speed in a MIDI sequencer program. But my main instrument is the guitar. Where most synthesizers include a keyboard to play the musical notes, I wanted to be able to use my electric guitar. Guitar synths have been available for many years, but most require you to purchase a special guitar having separate pickups for each string. I decided to build one final synthesizer that I could play using my Fender Telecaster instead of a keyboard.

Reading the frequency of notes played by a string instrument reliably is the "holy grail" of guitar synthesizers. Unlike a static wave containing harmonics, such as square and sawtooth waves, the harmonics of a guitar or bass string vary in frequency as the string stretches and relaxes at the fundamental frequency. This creates "rolling harmonics" that ride up and down the fundamental when viewed on an oscilloscope. **Figure 1** shows a

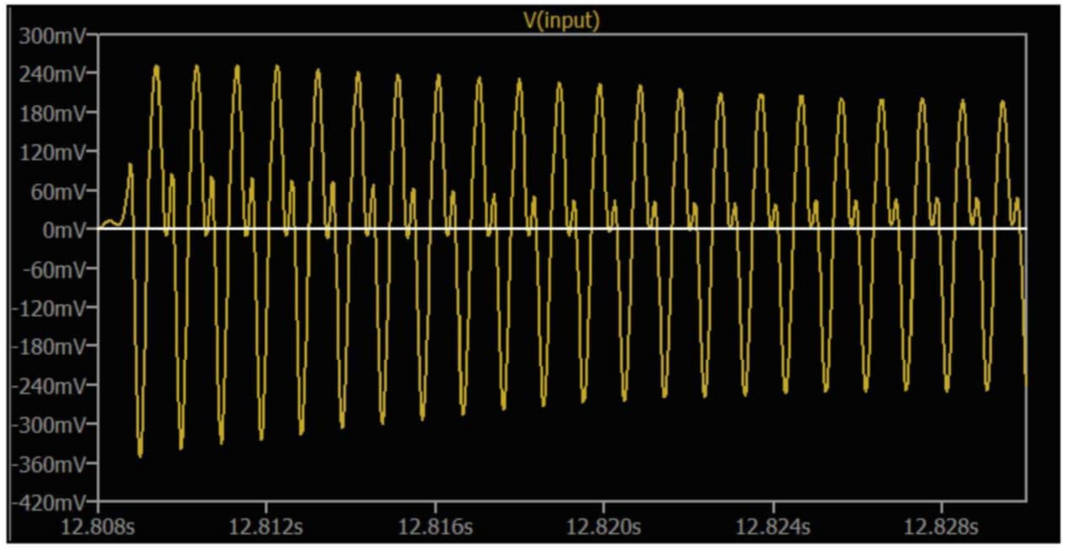


Figure 1: This high note shows the "rolling harmonics" that occur with all string instruments. Lower notes are even more complex, having many harmonic transitions through zero within each fundamental cycle.

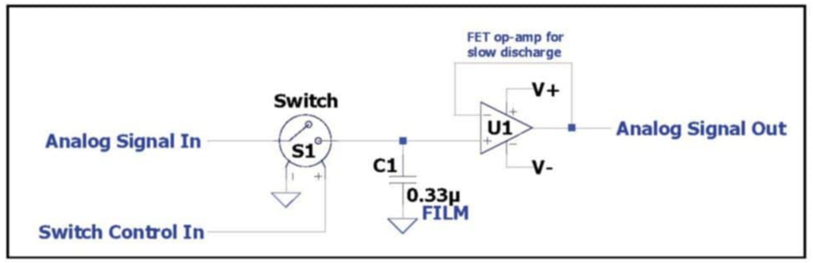


Figure 2: A basic Sample & Hold circuit is made from a voltage-controlled switch, a capacitor, and a buffer.

note played high up on the guitar neck. Between the large fundamental cycles you can see smaller third harmonics whose level and position on the main wave change as the note sustains. So at the start of the note there are two extra zero crossings for each cycle, but 20ms later (at the right edge) there's now only one zero crossing per cycle.

One common way to convert a frequency to an equivalent voltage is to generate a pulse having a fixed duration each time the input signal ascends through zero. These pulses are then accumulated and averaged by a low-pass filter. The higher the frequency, the more pulses that are accumulated, which creates a higher output voltage. But if the input wave passes through zero additional times due to the harmonics, the output voltage jumps one or more octaves higher than it should. So the first thing needed to read a guitar's frequency is a filter that removes the harmonics. This is surprisingly more difficult to do than you might think, requiring a very steep filter!

My guitar synthesizer first filters out the harmonics, then an equivalent voltage is generated with a Period-to-Voltage converter (P2V) that measures the length (period) of each cycle. This has one big advantage over the more common Frequency-to-Voltage (F2V) converter described earlier. Instead of having to accumulate and average pulses over several cycles, a P2V converter assesses the time span between zero crossings. So it can determine the input frequency on the first cycle!

That voltage is then sent to Voltage-to- Period (V2P) oscillators that generate equivalent frequencies. When the note ends and the input voltage subsides, a noise gate senses the lack of signal and shuts off the guitar input. Then the last voltage that was measured can be held, similar to the Sustain pedal on a piano. Unfortunately, by the time the noise gate closes, the last voltage saved is invalid because the final bits of the note are mostly noise.

The Sample & Hold Time Machine

A Sample & Hold (S&H) circuit is very simple comprising a switch, a capacitor, and an op-amp or other "buffer" to prevent draining the voltage currently stored in the capacitor. **Figure 2** shows a basic Sample & Hold circuit. The analog input passes through a voltage-controlled switch, which then charges a capacitor. Whenever the switch closes, the current input voltage is "sampled" and applied to the capacitor. Then when the switch opens, the capacitor's present voltage continues to be sent to the output regardless of the input. That's the "hold" part. An FET op-amp is often used because its input bias current is too small to drain the charge held in the capacitor.

To avoid sustaining the corrupt voltage that happens at the end of a guitar note requires looking back at least 20mS before the note ended, then quickly switch to sending that to the V2P oscillators. But it's not enough to simply grab a backup voltage sample every 20mS because the previous sample might have been taken just before the current note ended. To do this reliably with analog electronics, a round-robin set of three S&H capacitors is needed to save the three most recent voltages. While a note plays, a single-pole three-throw (SP3T) electronic switch alternates sending the incoming P2V voltage stream to each S&H in turn. This makes available for output to the oscillators: 1) the current stream; 2) the sample taken and held less than 20mS earlier; or 3) the sample taken and held 20mS before that.

LTspice

Designing a large complex system, such as an entire synthesizer, is expensive and time consuming, buying and connecting hundreds of components. The freeware LTspice program from IC maker Analog Devices provides a fully stocked parts bin with a complete set of test gear. You just draw the schematic, then "run" it to see if it works. I used LTspice for my October 2017 Mojo Maestro article to determine the best resistor and capacitor values for that simple circuit. LTspice is a deep program with a modest learning curve. Thankfully there's a very active support group that answers questions and offers advice. One huge feature of LTspice is it can accept input from a .wav file, and send output to a .wav file. This really helps when designing audio circuits!

Referring to the full S&H circuit in **Figure 3**, NE555 timer IC (U5) serves as an oscillator running at 50Hz (20 mS). This clocks a Modulo 3 Counter to generate the binary address sequence 0, 1, 2, 0, 1, 2 ... which directs Input Switch #1 (U3) to feed each S&H cap in turn. This switch has two address

Digital Logic in LTspice

Logic ICs are used to make logical decisions in both analog and digital circuits. For example, if a voltage comparator has a positive output and a timer IC is still active, an AND gate will output a "1" (positive voltage). An OR gate is similar, outputting a 1 if any of its inputs—1 OR 2 OR 3—are positive. There are thousands of logic ICs available, so rather than provide all those component models LTspice offers generic versions that are highly flexible. The AND and OR gates each have five inputs, with both normal 1 and reversed 0 outputs. So the same AND gate can also serve as a NAND (Not AND) gate. In **Figure 3**, the outputs with the little circle are the "Not" output connections. When more than five inputs are needed, multiple gates can be daisy-chained together.

There isn't room to describe all of the logic devices LTspice provides, but I'll mention the D Flip-Flops A2 and A3 since they're used in the Sample & Hold circuit. Flip-Flop means the output goes back and forth between the two states 1 and 0, and the "D" stands for Data. This component is often called a D-Flop for brevity. When a 1 arrives at the Clock input, whichever value—0 or 1—is present at the Data input is sent to the Q (normal) output. The "Q-bar" output is always the opposite polarity of the Q input, much like the Not output of the AND and OR gates. So one use for a D-Flop is to save one bit of memory. You put a 1 at its input, send a clock pulse, and that 1 is remembered until the next clock pulse. However, D-Flops also have Preset and Clear inputs, to force the Q output to 1 or 0, respectively, regardless of the Data and Clock inputs.

With all of the LTspice digital components, any inputs or outputs that are not used should be grounded. This tells LTspice not to waste time computing those connections. Of course, in a real circuit you'd never connect an unused output to ground! For mixed analog and digital circuits like this Sample & Hold, you also need to specify the input and output operating voltages. Most logic chips work at either 3V or 5V, but this circuit uses 15V power supplies. If you download the project file, right-click on any of the digital parts and you'll see two user-defined values: *Vhigh 15* and *tD 10n*. The first sets the operating voltage to 15, common with CMOS logic chips, and the second creates a propagation delay of 10 nanoseconds through the component. Most digital logic circuits work correctly with no delay, but some circuits cross-connect outputs to inputs and can lock up without this delay. So it's good practice to always add the delay.

About the Author

Ethan Winer has been an audio engineer and professional musician for more than 45 years. His Cello Rondo music video has received nearly 2 million views on YouTube and other websites, and his book *The Audio Expert* published by Focal Press, now in its second edition, is available at amazon.com and Ethan's own website ethanwiner.com. Ethan is also a principal at RealTraps, a US manufacturer of high-quality acoustic treatment.

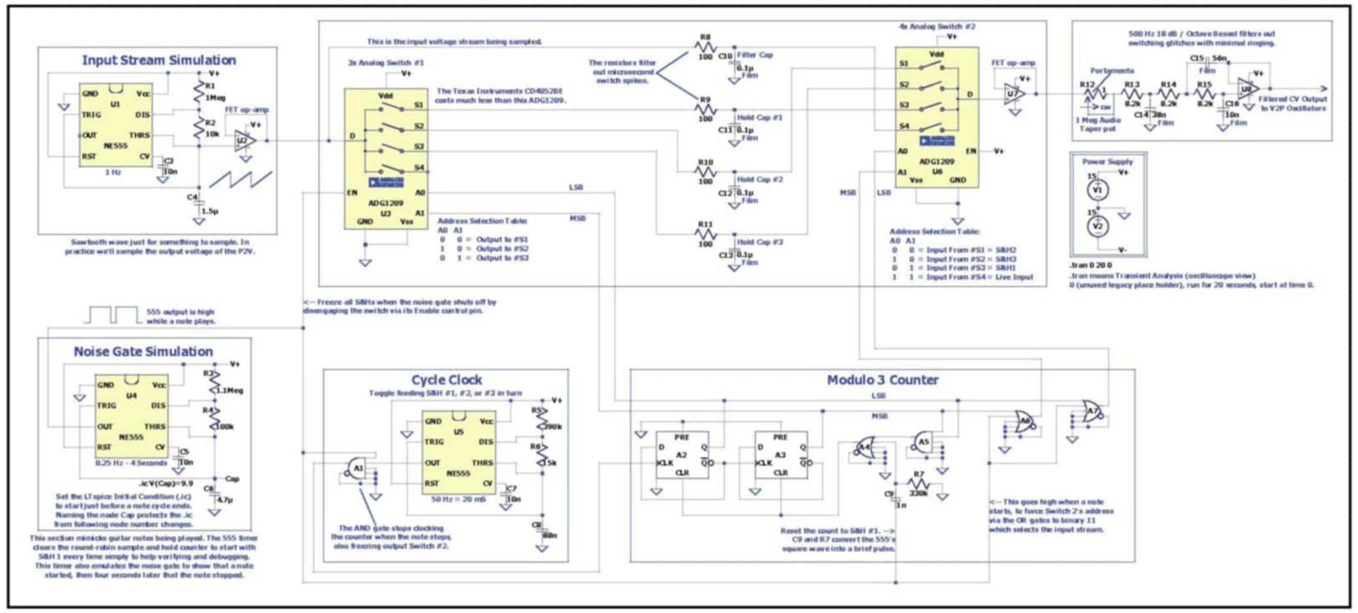


Figure 3: I call this section of the guitar synthesizer the Time Machine because it lets you go back in time to retrieve an input voltage that no longer exists. High-quality PDF schematic available online. Visit the audioXpress Supplementary Material website page.

lines telling it which input should go to its output. In logic speak, LSB stands for Least Significant Bit and MSB is Most Significant Bit. Dedicated counter ICs are available, but it's just as easy to create a counter using a pair of D Flip-Flops.

Normally this counter would count from 0 to 3 instead of 0 to 2, so AND gate A5 resets the count back to 0 when both outputs go high (binary 11 = 3) via the D-Flop Clear inputs. OR gate A4 is in the reset path simply to allow also clearing at the start of each note. Rather than belabor the more subtle details of this circuit here, comments on the schematic explain more about how each section works.

When a note starts and the noise gate opens, SP4T Output Switch #2 (U6) selects the incoming P2V voltage stream as its output to the V2P oscillators via OR gates A6 and A7. When the note stops, the most recently held sample is likely less than 20mS old, so we need to grab the sample that was captured and held before that one. One method subtracts 2 from the Counter address of Input Switch #1, then uses the result to tell Output Switch #2 which S&H was two counts earlier. But it's even easier to use the same input switch address for the output switch, but swap around the output switch's inputs. So if S&H #3 is active when the note ends, S&H #1 is selected for the output because it's connected to the output switch's input #3.

The incoming P2V voltage stream is not yet filtered to average out the occasional errant cycle that could cause an audible glitch. So after choosing between S&H #1 through #3, Output Switch #2 feeds the Portamento control (set to 1Ω in the schematic to disable it), then a three-pole low-pass filter. This filter applies minimal averaging to the datastream, and also removes any brief spikes caused by capacitive leakage between the digital and analog paths within Switch #2. Likewise, simple RC filters are placed after the input switch for the same reason, though they're not strictly needed. The ADG1209 and CD4052BE switches referenced in the Figure 3 schematic have an internal On resistance of around 120Ω. So R8 through R11 simply add a little more resistance. I used the Analog Devices ADG1209 switch because I couldn't find a component model for the less expensive CD4052BE.

Author's Note: The schematics for this article are in PDF files you can download from the Project Files link. This lets you zoom in to read them more clearly.

Project Files

To download additional material and files, including the Sample & Hold LTspice file, visit http://audioxpress.com/page/audioXpress-Supplementary-Material.html

Resources

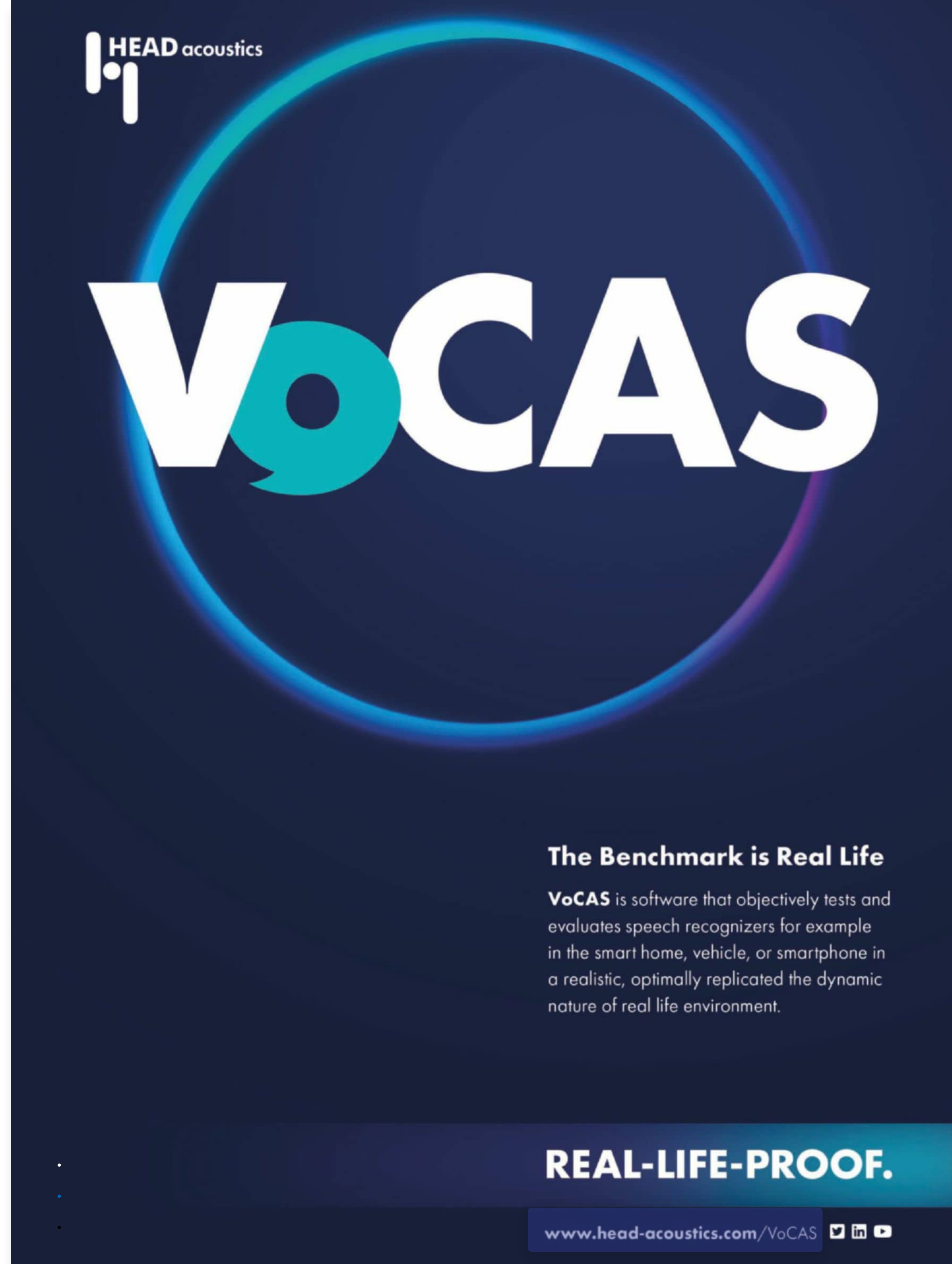
HP computer synthesizer tune: http://ethanwiner.com/phils-hp.mp3

LTspice: www.analog.com/en/design-center/design-tools-and-calculators/ltspice-simulator.html

LTspice online support forum: https://groups.io/g/LTspice

E. Winer, "Build the Mojo Maestro," *audioXpress*, October 2017, https://audioxpress.com/article/you-can-diy-build-the-mojo-maestro

17572091



Amplifier Comparisons Revisited

By
Richard Honeycutt

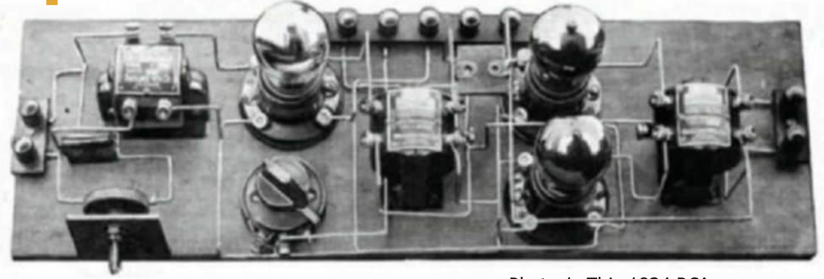


Photo 1: This 1924 RCA Class-B push-pull amplifier was a true trendsetter.

After years of research and much debate, our columnist has determined that the superiority of tube or solid-state amplifiers is truly a matter of perception. He believes that those people who are convinced that tube amps or solid-state amps are better because they can hear a difference do indeed perceive a difference; however, it is not necessarily a measurable or objective difference.

Longtime readers of *audioXpress* might remember Hollow-State Electronics articles from the August through November issues of 2012 that discussed the so-called "tube sound." In fact, disagreements about the "sound" of different audio amplifiers have persisted since the early days of audio electronics. The first electronic audio amplifiers were single-ended Class-A devices, whose output power was limited by the inherent inefficiency of that circuitry. In 1915, Edwin H. Colpitts of Bell Telephone Laboratories was granted a patent that included a drawing of a push-pull amplifier, although the push-pull principle had previously been claimed in an 1895 patent by W. W. Dean. (The Dean patent predated the era of electronic amplifiers.) RCA pioneered the use of Class-B push-pull power amplifiers.

Benefits of Class-B Amplifiers

Advocates of Class-B push-pull amplifiers claimed two primary advantages: (1) the maximum audio power output from a push-pull tube pair was over three times that of a pair of Class-A-connected tubes; and (2) the push-pull (or "balanced") connection cancelled even-order harmonic distortion.

The 1924 RCA Radiola III regenerative radio receiver incorporated the "balanced amplifier," a Class-B push-pull design (**Photo 1**). Home broadcast radio receivers comprised most of the market for audio electronics in the early 1920s, but the most critical arena for audio quality was the cinema. Western Electric and RCA competed for dominance in that market. Audiophiles argued vigorously about the relative merits of single-ended versus push-pull amplifiers. The primary disadvantage of early push-pull amplifiers is that some of them used Class-B operation, and crossover distortion was audible (**Figure 1**).

With true Class-B biasing, each tube conducts for exactly one half-cycle. Crossover distortion represents the failure of the output wave to transition perfectly from one tube to the other, and comprises strong higher-order harmonic distortion products that are very annoying. Class-AB biasing, in which both output tubes conduct at crossover, greatly diminishes crossover distortion. Although some purists still maintain that Class-A single-ended tube power amps sound better than push-pull ones, push-pull amps have been the de facto standards for many years.

Frequency Response

During the 1940s and the 1950s, much of the discussion about amp sound was focused on the frequency response. With the advent of "high-fidelity" sound systems, most manufacturers published frequency response specs that included the low- and high-frequency limits and a response tolerance (+X dB). The accuracy of these published specs varied widely. The same can be said for published total harmonic distortion (THD) and intermodulation distortion (IMD) specs. Other measurable performance metrics (e.g., damping factor) came under scrutiny.

With the introduction of solid-state audio amps in the 1960s, crossover distortion reared its ugly head again as some years were required before true complementary-symmetry designs were developed, and stable biasing arrangements for these new solid-state amps. Then in the late 1960s, Finnish Researcher Matti Otala discovered transient intermodulation distortion (TIM), which resulted from amplifiers having insufficient slew rates, and created brief "buzz bursts" similar to those resulting from Gibbs' phenomenon.

The combination of wide frequency response and low slew rate (maximum output change in volts/microsecond) appeared primarily in solid-state amps, so Otala had indeed identified a mechanism that would explain a perceived difference in the sound of solid-and hollow-state amps. Within a few years of Otala's publication of a low-TIM amp design [1], also called Slewing-Induced Distortion (SID), TIM was pretty much history.

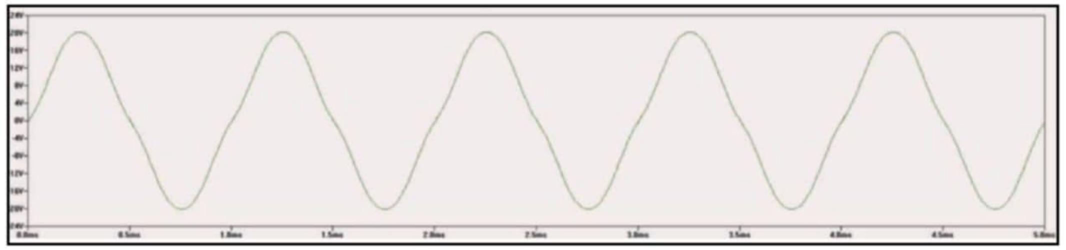


Figure 1: The "crossover notch" where the output tubes' conduction changes from one to the other is an annoying form of distortion.

Tube vs. Solid State

The difference in opinion concerning the superiority of tube or solid-state amplifiers has been dubbed "The Great Debate." In today's audio market, not only is the hollow-state amplifier genre an important player, but new releases of old classic amps are garnering a lot of interest. **Photo 2** and **Photo 3** show a recent design and a redesign of a, or some would say, "the," classic tube amp. Purchasers' choices of sides in "The Great Debate" can be categorized in four groups:

- Those who believe tube amps are inherently better because they are convinced that they can hear a difference.
- Those who believe that all "serious" audiophiles know that tube amps are better, so that must be the case.
- Those who prefer solid-state amps and are convinced that they can hear a difference.
- Those who choose not to take a side.

Photo 2: The Vacuum Tube Logic ST-85 stereo amplifier uses four EL34 tubes and delivers 60W per channel.

Within the first group, there are two subgroups: those who believe that there is a measurable performance factor that accounts for the difference, and those who believe that whatever the factor is, we have not yet identified what it is, or perhaps we don't know how to measure it.

Efforts have been underway for decades to settle "The Great Debate." The most productive efforts have begun by trying to determine whether listeners can in fact hear a difference among different well-designed amps. Experiments by Dr. John Vanderkooy of the Audio Research Group at the University of Waterloo and Dr. Floyd Toole, Corporate Vice President—Acoustical Engineering at Harman International from 1991 until he retired in 2007, and Senior Research Officer in the Acoustics and Signal Processing Group at the National Research Council of Canada—have revealed that we humans are particularly susceptible to self-deception in judgment of audio quality. Listeners who were asked to choose the better loudspeaker from a pair commonly chose the most highly regarded brand, even when the drivers and crossovers had been swapped, so that, for example, the "JBL" speaker was in fact a JBL cabinet containing Audio Research components.

Photo 3: The Western Electric 91E integrated amp is now available for pre-order.

Photo 4: The Evenstar Pro Peregrin studio monitor was used in recording the test files.

Test Protocols

From these experiments, robust protocols have been developed to eliminate as much unconscious listener bias as possible. These include careful level matching of the outputs of the amps being compared (audio as little as 1dB louder is generally rated by most listeners as sounding better). Extreme care in frequency response matching has been used in some tests. Since psychological and parapsychological experiments ever since the 1930s have indicated the ability of test administrators to somehow unintentionally transmit their own biases to test subjects, double-blind testing (neither the administrator nor the subject knows which amp is which) has been used. A particularly effective protocol is A/B/X testing.

In October 1991, David Clark of DLC Design presented the paper, "Ten Years of A/B/X Testing," at the 91st Convention of the Audio Engineering Society (AES). (A/B/X amplifier tests compare an unknown amplifier "X" with two known amplifiers, "A" and "B".) The test subject's goal is to determine whether "X" is "A" or "B". For example, an excellent amplifier can be used as "A", and a medium-grade amplifier as "B". The test subject can switch at will among "A", "B", and "X", but the switches may be connected so that position "X" is actually amplifier "A". If the subject can reliably identify that "A" is "X", then clearly the difference between "A" and "B" is perceptible to him/her. If not, we cannot conclude that there are physical causes for the differences that some listeners hear under less-controlled conditions.

A/B/X tests are usually double-blind, but do not require that the brand/model of the equipment under test be kept from the listeners, since neither the listener nor the test operator knows which is "A", "B", or "X". Test subjects are not allowed to communicate with each other.

In an A/B/X test, the listener chooses when to flip the switch, allowing whatever amount of time (s)he feels is needed to properly identify the unit under test as "A" or "B". A test may span several listening sessions, if the listener so chooses, or may be finished quickly if the listener is confident (s)he has determined the identity of "A" or "B" as "X" in a short time.

A/B/X testing excels at finding perceptible differences, if any exist, but is not designed to establish levels of accuracy or preferability. The A/B/X test itself was compared with long-term listening as a method of identifying a calibrated 2.5% THD component that was added to a musical signal. The Audiophile Society acted as the "golden ears;" and the Southwestern Michigan Woofer and Tweeter Marching Society (SMWTMS) acted as the "engineers."

Neither group could identify the distortion at a 5% confidence level in long-term listening tests. However, using A/B/X testing, the SMWTMS not only correctly proved the audibility of the distortion in 45 minutes of testing, but also correctly identified a lower amount of distortion. In the complete series of tests, THD was found to be audible at 4% using big band jazz music, 2% on flute music, and 0.4% on a sine wave. The spectrum of the harmonic distortion was not specified.

Clark added a note in his 1991 paper, based upon a private conversation with Thomlinson Holman (the "TH" of THX®). Holman has found that a number of professional power amplifiers do distort audibly when driving highly reactive loads (e.g., some theater speakers when playing explosive movie sounds). The cause could well be that under such severe load conditions, the amps' power supplies experience instantaneous drops in voltage. This condition would be easy to identify using an oscilloscope.

Author's Opinion

Following a detailed discussion of "tube sound" with the late Edward T. Dell, founder of *The Audio Amateur* (the predecessor magazine of

About the Author

Dr. Richard Honeycutt bought a *Popular Electronics* magazine from the rack at a bus station in 1960. It contained an article on building a transistorized audio amplifier, which captured his interest. He followed up by subscribing to several electronics magazines. To avoid the cost of parts, he built up a "junk box" by disassembling trade-in TVs from his uncle's furniture store. His dad bought him the Van Valkenburg, Neville, and Nooger Basic Electronics book set, so he learned about tube circuits as well as transistors. He started repairing electronic devices about 1965, earning his First-Class Commercial FCC license in 1969. He has worked part-time in broadcast radio engineering and audio electronics repair since 1968, in addition to full-time work in acoustical and audio system design, plus a 20-year stint teaching electronics at the college level.

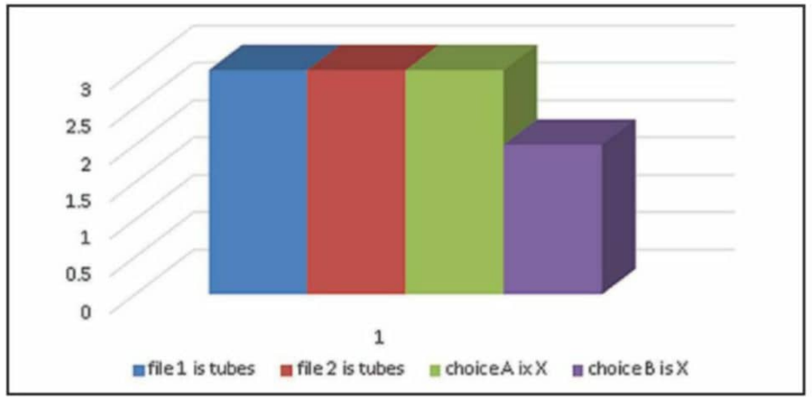


Figure 2: The small number of respondents did not agree on their judgments in the amp test.

audioXpress). I adopted his opinion that those who are convinced that tube amps or solid-state amps are better because they can hear a difference do indeed perceive a difference. It is not necessarily a measurable or objective difference (one that other people can hear). The consistent failure of well-designed experiments to show that listeners can identify an aural difference among different amps tested with careful frequency response and level matching and avoidance of overdriving (a difficult condition to enforce) leads me to the conclusion that there is no one superior amp design that will sound best to every listener in every application.

Online Test

In the November 2012 Hollow-State Electronics article, I announced an online listening test. The test had two parts. In the first, participants were invited to listen on quality headphones (not computer speakers) to a pair of .wav files of a short clip played by the lead guitarist of Darrell Harwood and the Cold Creek Band. One file was made by recording direct to digital from the guitar, then playing the recording through a solid-state power amplifier adjusted to produce an acceptable amount of distortion for proper artistic effect. The other file used the same digital recording, played through a tube amplifier adjusted in the same way.

Both recordings were made using the same speaker and cabinet, and were recorded using a type 1 calibrated measurement microphone with a tensioned stainless steel diaphragm. The same physical setup was used for both recordings, and the levels of the recordings were carefully matched. Listeners could play the files as many times as desired, and were then requested to send an email stating which file, "A" or "B", sounded better. If they chose to include an opinion as to which was the tube-amp recording, and which was solid-state, they could do that as well. The purpose of this part of the test was to illustrate the sound of the different distortion spectra of tube vs solid-state amplifiers.

The second part of the test consisted of three .wav recordings of a brief clip of classical music. There was a reference clip made directly from the CD, a clip of the selection being played through a McIntosh hollow-state power amplifier, and a clip of the selection being played through a commercial solid-state power amplifier. Both of the "amplifier" recordings were made using an Evenstar Pro Peregrin studio monitor speaker (**Photo 4**), recorded using the measurement mic previously described, with the same physical setup. No equalization was applied to either amplifier, and the amplifiers were operated well below clipping. The reference ("X") straight-through recording was identified, and participants were asked to listen to the "A" and "B" recordings, determine which sounded more like the reference, and e-mail their results. The test was announced as being available through the end of 2012.

All told there were only eight participants in the online test. This number is far too small to draw any statistically meaningful conclusions. **Figure 2** shows the results just for the readers' interest. In addition to the graphed results, there were two participants who could not tell any difference between "A" and "B" in the A/B/X test, and two who said they liked guitar recording 2 better, but had no idea which was the tube amp.

Although the amp test was inconclusive in terms of establishing an audible difference between tube and transistor amps, it is useful to note that even in this small test, the reported perceived differences were very small. Yet some participants expressed strong opinions about the differences they heard. Certainly the differences are real, but we have not been able to establish that they are objective differences: different listeners perceive them in different ways, if at all.

Editor's Note: audioXpress *articles from 2001 to present can be found on the aX Cache, a USB drive available from www.cc-webshop.com.*

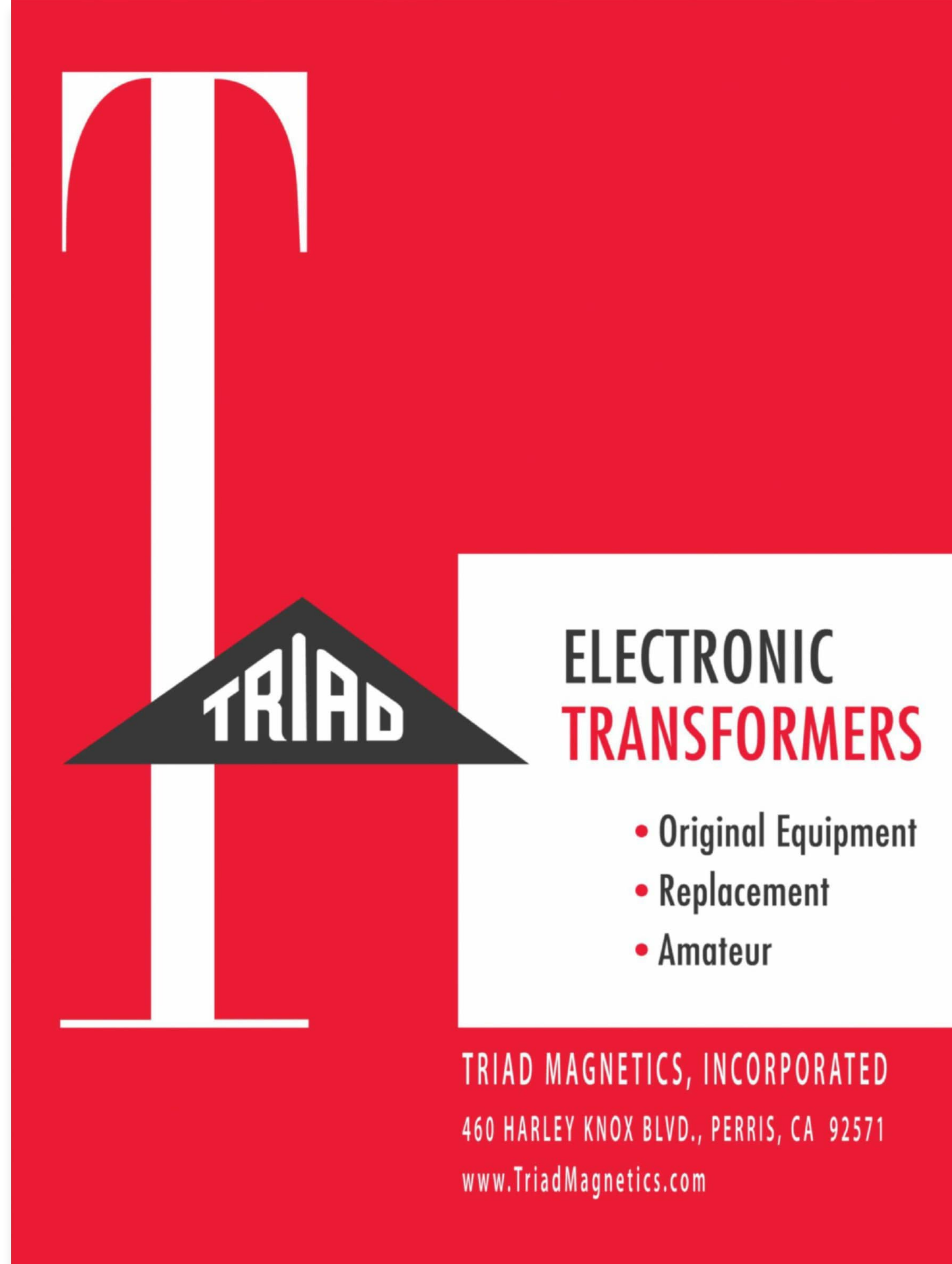
References

[1] M. Otala, "Circuit Design Modifications for Minimizing Transient Intermodulation Distortion in Audio Amplifiers," *Journal of the Audio Engineering Society*, Volume 20, Section 5, pp. 396-399; June 1972.

Resources

D. Clark, "Ten Years of A/B/X Testing," presented at the 91st Convention of the Audio Engineering Society, October 1991.

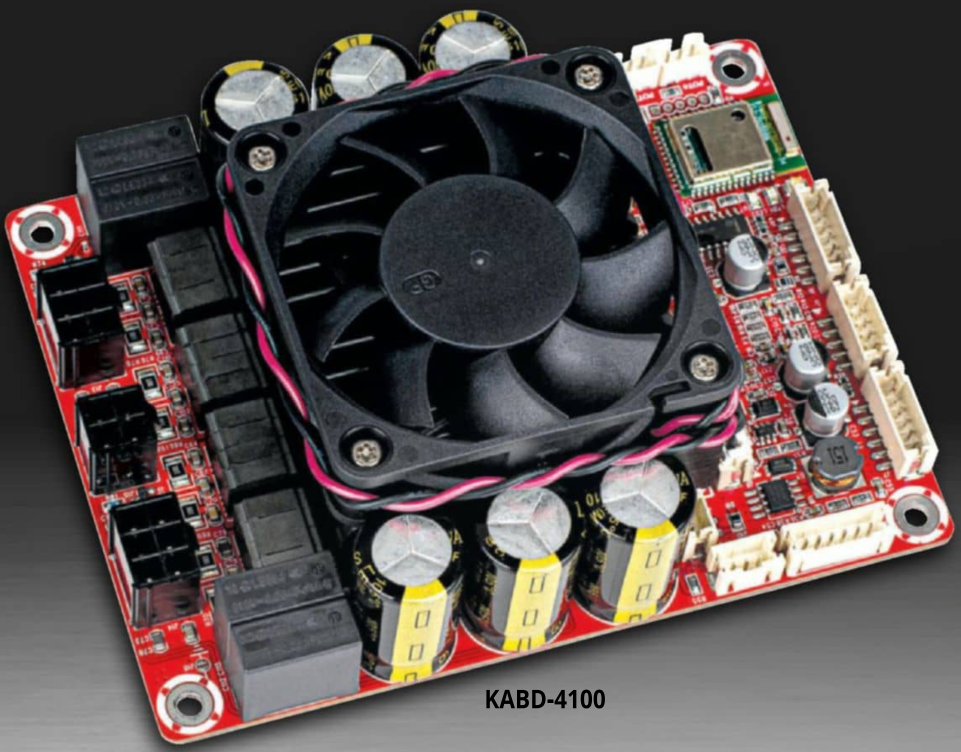
R. Honeycutt, "Differences in Amp Sound: What's the Truth?," *audioXpress*, November 2012.



New KABD DSP Amplifier Boards from Dayton Audio

THE NEXT LEVEL OF DSP AMPLIFICATION

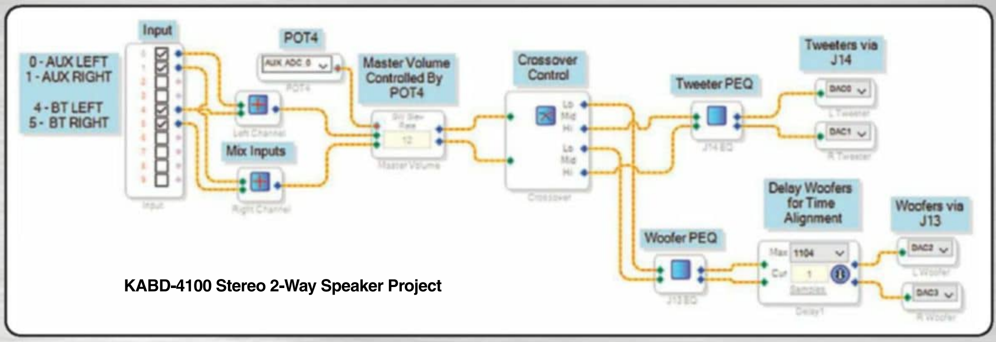
The Dayton Audio KABD line of amplifier boards give you unprecedented control for your next boombox project, active bookshelf speakers with DSP crossovers, 2.1 desktop system, or anything you can imagine. Featuring integrated digital signal processing, you can control your sound signature using free SigmaStudio software. With easy to configure solderless connections, drag and drop DSP, and predesigned examples, you are just clicks away from your own unique sound.



KABD-4100

What's New:

- All new versions: 2 x 30W, 2 x 50W, 4 x 30W, 4 x 100W
- Bluetooth 5.0 with aptX HD for high quality and long-range streaming
- Optimized standby mode and efficient amplification is perfect for battery-powered projects
- External potentiometer ports allow easy integration of volume, tone control, filters, and more for any project



KABD-4100 Stereo 2-Way Speaker Project

DAYTONAUDIO

daytonaudio.com

Distributed by:

PARTS EXPRESS

AUDIO COMPONENTS & SOLUTIONS

parts-express.com 1-800-338-0531

Custom OEM inquiries welcome

Scan here to learn more about the Dayton Audio KABD Amp Boards